PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/10, 15/25, 15/62, C07K 14/535, 14/005 | A1

(11) International Publication Number: **WO 00/18905**

(43) International Publication Date: 6 April 2000 (06.04.00)

(21) International Application Number: PCT/US99/20891

(22) International Filing Date: 24 September 1999 (24.09.99)

(30) Priority Data:
60/101,908 25 September 1998 (25.09.98) US

(71) Applicant *(for all designated States except US)*: G.D. SEARLE & CO. [US/US]; Corporate Patent Department, P.O. Box 5110, Chicago, IL 60680–5110 (US).

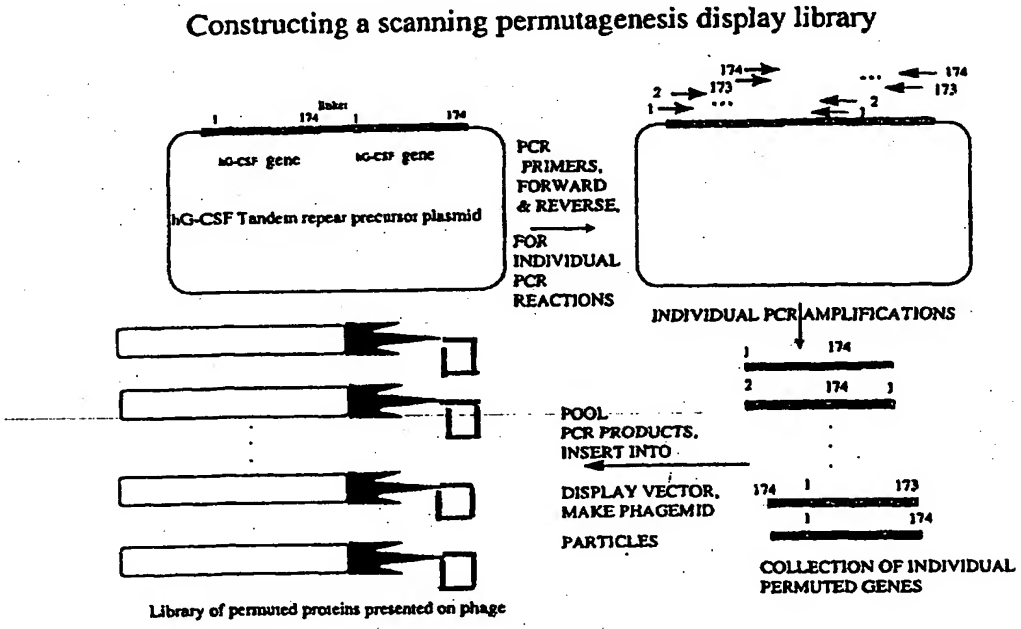
(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: LEE, Stephen, C. [US/US]; 828 Fernview Drive, Creve Coeur, MO 63141 (US).

(74) Agents: BAUER, S., Christopher et al.; G.D. Searle & Co., Corporate Patent Dept., P.O. Box 5110, Chicago, IL 60680–5110 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: METHOD OF PRODUCING PERMUTEINS BY SCANNING PERMUTAGENESIS



(57) Abstract

A method of producing circularly–permuted proteins (permuteins) by scanning permutagenesis comprises making and inserting a series of circularly–permuted genes into a display vector, expressing these genes such that the gene products are localized to the surface of the display vector, generating a library of display vectors presenting the permuted protein, affinity–selecting the display vectors with a target protein that can bind the permuted protein, isolating and analyzing clones of selected display vectors to identify the circularly–permuted protein. The invention further discloses methods of expressing and uses of permuteins.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

Method of producing permuteins by scanning permutagenesis

Priority

The present application claims priority under Title 35, United States Code, § 119 of United States Provisional Application Serial No. 60/101,908, filed September 25, 1998.

Field of the invention

A method of producing circularly-permuted proteins (permuteins) by scanning permutagenesis comprises making and inserting a series of circularly-permuted genes into a display vector, expressing these genes such that the gene products are localized to the surface of the display vector, generating a library of display vectors presenting the permuted protein, affinity-selecting the display vectors with a target protein that can bind the permuted protein, isolating and analyzing clones of selected display vectors to identify the circularly-permuted protein. The invention further discloses methods of expressing and uses of permuteins.

Background of the invention

Protein permutagenesis

Circularly permuted proteins are made by reordering the primary sequence of a parent protein. The amino and carboxy terminal ends of the parent protein are joined by a peptide linker and new amino and carboxy terminal ends are generated at other positions in the sequence. This technique of generating variants has been applied to a wide variety of proteins (Table 1).

Circularly permuted proteins, in many cases, are structurally and functionally similar to their non-permuted parent molecule after they undergo refolding. The information necessary to direct the folding of proteins into tertiary structures is present in secondary structural domains. Vectorial folding of proteins from their native amino to carboxy ends is not often observed. The ability of permuteins to retain structural and functional properties is remarkable. extending earlier observations on the plasticity of proteins with respect to amino acid

substitutions (Olins P.O. et al., *J. Biol. Chem.* **270**: 23754-23760, 1995; Lowman and Wells, *J. Mol. Biol.* **234**: 564-578, 1993) and short amino acid insertions (Sondek, J. and D. Shortle, *Proteins* **7**: 387-393, 1990; Shortle, D. and J. Sondek, *Curr. Opin. Biotechnol.* **6**: 299-305).

Protein sequence reorganization

Rearrangements of DNA sequences serve an important role in evolution by generating a diversity of new proteins differing in structure and function. Gene duplication and exon shuffling, for example, generate diversity and provide organisms with a competitive advantage since the basal mutation rate is low (Doolittle, *Protein Science* **1**: 191-200, 1992).

Recombinant DNA methods have facilitated studies on the effect of sequence transposition on protein folding, structure, and function. The first rearrangement of proteins using this approach was described by Goldenberg and Creighton (*J. Mol. Biol.* **165**:407-413, 1983). A new N-terminus is selected at an internal site (breakpoint) of the original sequence, the new sequence having the same order of amino acids as the original from the breakpoint until it reaches an amino acid that is at or near the original C-terminus. At this point the new sequence is joined, either directly or through an additional portion of sequence (linker), to an amino acid that is at or near the original N-terminus, and the new sequence continues with the same sequence as the original until it reaches a point that is at or near the amino acid that was N-terminal to the breakpoint site of the original sequence, this residue forming the new C-terminus of the chain. Similar approaches have also been used in other studies (Cunningham et al., *Proc. Natl. Acad. Sci. U.S.A.* **76**:3218-3222, 1979; Teather & Erfle, *J. Bacteriol.* **172**: 3837-3841, 1990; Schimming et al., *Eur. J. Biochem.* **204**: 13-19, 1992; Yamiuchi and Minamikawa, *FEBS Lett.* **260**:127-130, 1991: MacGregor et al., *FEBS Lett.* **378**:263-266, 1996).

These general approaches have been applied to proteins which range in size from 58 to 462 amino acids (Goldenberg & Creighton, *J. Mol. Biol.* **165**:407-413, 1983; Li & Coffino, *Mol. Cell. Biol.* **13**:2377-2383, 1993). The proteins represent a broad range of structural classes, including proteins that contain predominantly alpha helix (interleukin-4; Kreitman et al., *Cytokine* **7**:311-318, 1995), beta sheet (interleukin-1; Horlick et al., *Protein Eng.* **5**:427-431, 1992), or mixtures of the two types of secondary structures (yeast phosphoribosyl anthranilate isomerase; Luger et al., *Science* **243**:206-210, 1989).

Although broad categories of protein function are represented in these sequence reorganization studies, the results of these studies have been highly variable. In many cases substantially lower activity, solubility, or thermodynamic stability were observed (*E. coli* dihydrofolate reductase, aspartate transcarbamoylase, phosphoribosyl anthranilate isomerase, glyceraldehyde-3-phosphate dehydrogenase, ornithine decarboxylase, ompA, yeast phosphoglycerate dehydrogenase). In other cases, the sequence rearranged protein appeared to have many nearly identical properties as its natural counterpart (basic pancreatic trypsin inhibitor, T4 lysozyme, ribonuclease T1, Bacillus β-glucanase, interleukin-1β, α-spectrin SH3 domain, pepsinogen, interleukin-4). In exceptional cases, an unexpected improvement over some properties of the natural sequence was observed, e.g., the solubility and refolding rate for rearranged α-spectrin SH3 domain sequences, and the receptor affinity and anti-tumor activity of transposed interleukin-4-*Pseudomonas* exotoxin fusion molecule (Kreitman et al., *Proc. Natl. Acad. Sci. U.S.A.* **91**:6889-6893, 1994; Kreitman et al., *Cancer Res.* **55**:3357-3363, 1995).

The primary motivation for reorganization studies has been to study the role of short-range and long-range interactions in protein folding and stability. Sequence rearrangements of this type convert a subset of interactions that are long-range in the original sequence into short-range interactions in the new sequence, and vice versa. The fact that many of these sequence rearrangements are able to attain a conformation with at least some activity is persuasive evidence that protein folding occurs by multiple folding pathways (Viguera et al., *J. Mol. Biol.* **247**:670-681, 1995). In the case of the SH3 domain of alpha-spectrin, choosing new termini at locations that corresponded to beta hairpin turns resulted in proteins with slightly less stability, but which were nevertheless able to fold.

The positions of the internal breakpoints used in the studies cited above are found exclusively on the surface of proteins, and are distributed throughout the linear sequence without any obvious bias towards the ends or the middle (the variation in the relative distance from the original N-terminus to the breakpoint is ca. 10 to 80% of the total sequence length). The linkers connecting the original N- and C-termini in these studies have ranged from 0 to 9 residues. In one case (Yang & Schachman, *Proc. Natl. Acad. Sci. U.S.A.* **90**:11980-11984, 1993), a portion of sequence has been deleted from the original C-terminal segment, and the connection made from the truncated C-terminus to the original N-terminus. Flexible hydrophilic residues such as Gly and Ser are frequently used in the linkers. Viguera et al.(*J. Mol. Biol.* **247**:670-681, 1995) compared joining the

original N- and C- termini with 3- or 4-residue linkers; the 3-residue linker was less thermodynamically stable. Protasova et al. (*Protein Eng.* 7:1373-1377, 1994) used 3- or 5-residue linkers in connecting the original N-termini of *E. coli* dihydrofolate reductase; only the 3-residue linker produced protein in good yield.

Protein permutagenesis can be used to optimize the activity of fusion proteins or proteins conjugated to other molecules. A fusion between interleukin-4 (IL-4) and *Pseudomonas* exotoxin has been permuted resulting in a protein that has the first amino acid of the IL-4 domain at position 38 and the new carboxy end occurs at amino acid position 37 (Kreitman, R. J. et al., *Proc Natl Acad Sci USA* **91**: 6889-6893, 1994). The permuted fusion has increased affinity for the IL-4 receptor, increased cytotoxicity to IL-4 receptor bearing renal carcinoma cells, and increased anti-tumor activity in a murine model, compared to the non-permuted parent fusion protein (Kreitman, R. J. et al., *Proc Natl Acad Sci USA* **91**: 6889-6893, 1994; Kreitman, R. J. et al., *Cancer Res.* **55**:3357-3363, 1995; Puri, R. K. et al., *Cellular Immunol.* **171**: 80-86, 1996). Increased potency of the permuted molecule is believed to result from a reduction in steric interference between the IL-4 domain in the parent molecule and its receptor.

Steric hindrance is likely to be a concern for other chimeric proteins which interact with receptors through a relatively large area of their surface. The same issue also arises with bioconjugates, containing relatively small chemicals conjugated to proteins or other molecules in complex polymers (Rose, K. et al., *Molecular Immunology* **32**: 1031-1037, 1995).

Phage display methods

Display methods allow affinity selection of protein variants from a library of displayed proteins or peptides (Clackson, T. and J.A. Wells, *Tibtech* **12**: 173-184, 1994; Winter, G., *Drug Development Res.* **33**: 71-89, 1994). Many biological entities can be used in display methodologies (so-called "genetic packages" for presentation, including bacterial and eukaryotic cells, various eukaryotic and prokaryotic viruses, and spores), but the most commonly used vehicles used for display are filamentous bacteriophage, as used herein. We envision the possibility that a genetic package other than the particular phage used here could be used to present libraries of permuteins, and if so, constitute essentially the same invention.

Foreign proteins are presented on the surface of a phage particle, and the gene encoding the foreign protein is encapsulated in the virion. Because they are linked by the phage particle, affinity isolation of the presented protein also leads to affinity isolation of the corresponding genes. Extremely large libraries of phage presented proteins are constructed and affinity screened very rapidly. From the standpoint of how quickly mutant proteins can be made and screened for activity, phage display is the most efficient mutagenesis technique currently available.

Functional properties of permuteins

Permuteins can have improved biological properties by acting through several mechanisms. The permutein acting on the same type of cell as its parent molecule, may have increased binding, or other action, by virtue of increased avidity. Dimers or higher order multimers of these proteins with themselves or other chemical groups, including proteins, can have increased efficacy or potency, or both.

Permuteins can also have improved therapeutic properties through a variety of mechanisms such as: (1) alterations in the overall on- or off-rates or K_a or K_d of the ligand(s) on the target cell; (2) activation or blockade of complementary receptor signaling pathways; and/or (3) more specific targeting of to the cell of interest. The permuteins may also possess a unique pharmacokinetic distribution and clearance profile (Dehmer et al., *Circulation*, **91**, 2188-2194, 1995; Tanaka et al.,. *Nature Medicine*, **3**, 437-442, 1997).

Permuteins can also have improved properties *in vivo*, compared to the two components individually, as a result of alterations in biodistribution or half-life. The improved properties can also result from the binding of the permutein to one or more of the receptors, pharmacokinetics, or uptake of the permutein is altered in a favorable manner.

Molecular biology approaches have traditionally been used to permute proteins (Horlick, R.A. et al., *Protein Engineering* **5**: 427-431, 1992) although chemical approaches have been used to make small permuted proteins (Goldenberg, D. P. and T. E. Creighton, *J. Mol. Biol.* **165**: 407-413, 1983). These approaches are relatively labor intensive, limiting the number of permuteins that can be generated and efficiently screened for the desired biological activities. Rapid methods of generating permuteins, coupled with efficient methods for screening are needed that will result in the identification of novel active molecules.

Summary of the invention

The present invention is an improved method for generating permuteins (scanning permutagenesis) based on the display of proteins on bacteriophage surface proteins. Phage display is a powerful, yet convenient tool, traditionally used for mutagenesis and screening (Clackson, T. and J.A. Wells, *Tibtech* 12: 173-184, 1994). Improvements to this technology allow the rapid generation and screening of libraries of permuteins. Variables, such as position of the new termini and the length and composition of peptide linkers can easily be varied to generate libraries of the desired diversity.

The present invention relates to methods of producing biologically-active circularly permuted proteins of the formula C^1-L^1-N^1, derived from a parent protein of the formula N^1-C^1, wherein C^1 is comprised of a segment derived from the carboxy portion of said parent protein; N^1 is comprised of a segment derived from the amino terminal portion of said parent protein; and L^1 is a chemical bond or a linker, linking C^1 to the amino terminus of L^1 and carboxy terminus of L^1 to the amino terminus of N^1; comprising the steps of: (a) making a series of circularly-permuted genes; (b) inserting said circularly-permuted genes into a display vector; (c) expressing said circularly-permuted genes such that the proteins encoded by said genes are presented on the surface of the display vector; (d) generating a library of display vectors presenting the expressed circularly permuted protein; (e) affinity-select the presenting display vectors with a target protein that can bind a biologically-active circularly-permuted protein; (f) isolate and analyze clones of selected display vectors to identify the presented circularly-permuted protein.

Preferably the method of making a series of circularly-permuted genes is selected from the group consisting of making a tandemly-repeated intermediate, total synthesis of a synthetic gene, assembly of a gene from synthetic oligonucleotides, DNA amplification, and limited digestion of a circular intermediate.

Preferably, the display vector is selected from the group consisting of bacteriophage display vectors, bacteria, and baculovirus vectors. Even more preferably the presentation vector is a bacteriophage. Even more preferably, the presentation vector is bacteriophage M13. Most preferably, the presentation vector is a bacteriophage M13 gene III vector.

Preferably the method of making a series of circularly permuted genes is a method of making a tandem repeat intermediate. Even more preferably circularly permuted genes are amplified from the repeat by gene amplification.

Preferably the method of affinity selection comprises the steps consisting of (a) binding said presentation display vectors to a target protein; (b) eluting said display vectors; (c) amplifying said display vectors; and (d) biopanning a pool of said amplified display vectors.

Preferably, the length of C^1 in the permutein is longer than the length of C^1 in said parent protein. More preferably, the length of C^1 in the permutein is shorter than the length of C^1 in said parent protein. Most preferably, the length of C^1 in the permutein is the same length as the length of C^1 in said parent protein.

Preferably, the length of N^1 in the permutein is longer than the length of N^1 in said parent protein. More preferably, the length of N^1 in the permutein is shorter than the length of N^1 in said parent protein. Most preferably, the length of N^1 in the permutein is the same length as the length of N^1 in said parent protein.

The invention also contemplates circularly permuted proteins of the formula C^1-L^1-N^1 made by the method of scanning permutagenesis. Preferably, the DNA sequence encoding said linker L^1 is selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 368.

Preferably, the circularly-permuted protein is the G-CSF receptor agonist domain of a species of mylepoietin (MPO). MPO is one member of a family of novel dual cytokine receptor agonists (McKearn, J.P., Myelorestorative activities of synthokine and myelopoietin. *In* Proceedings of the 1996 IBC Conference on Therapeutic Applications of Cytokines, pp. 4.3.1-4.3.18, 1996) which are amenable to manipulation by phage display (Merlin, S. et al., *Applied Biochemistry and Biotechnology* 67: 15-29, 1997; Lee, S.C., Phage presentation for cytokine engineering. IBC's Second International Conference on Display Technologies, 1997).

Brief description of the figures

Figure 1. Schematic depiction of scanning permutagenesis

Plate A of Figure 1 shows the strategy to generate a scanning permutagenesis phage display library. A plasmid containing directly-repeated

tandem copies of the hG-CSF gene, for example, is constructed by standard methods. The tandem repeat plasmid is used as the template for PCR amplification of genes encoding permuted proteins. Each copy of the G-CSF gene is indicated in light gray (turquoise), and a DNA segment encoding a peptide linker is indicated in dark gray (red).

In individual PCR reactions, oligonucleotide primers that initiate PCR polymerization at the first nucleotide of a chosen codon of G-CSF, and directing polymerization to the end of the tandem construct specifying the carboxy end of the protein encoded on the template is annealed to the tandem template. Also, a second specific primer is also annealed to the template that initiates polymerization at the last nucleotide of the codon encoding the amino acid immediately preceding the codon where polymerization begins with first primer, and which directs polymerization in the opposite direction from that first primer. Amplification between these two primers produces a DNA segment encoding a permuted protein. For example, amplification between the primer indicated by a black arrow initiating at codon 2 and the primer indicated by the blue arrow and initiating at the codon before 2 (codon 1) produces an amplified gene encoding a permuted protein whose amino terminal residue is amino acid 2 of the native protein, and whose final amino acid is amino acid 1 of the native protein.

A linker peptide is present between the first and final amino acids of the parent protein (residues 1 and 174 in this example). A total of 174 individual amplifications would produce a complete collection of all permuted proteins of this example. More limited collections containing only a selected set of permuteins can be made, as well as more extensive collections made from multiple tandem template plasmids, each containing a different linker sequence between the first and last residues of the two directly repeated tandem gene sequences. The collection of amplified segments can then be inserted into a phagemid presentation vector by standard methods. Phagemid particles produced from these presentation constructs are the scanning permutagenesis phage display library.

Plate B of Figure 1 shows the affinity screening of a phage display library (See Clackson, T. and J.A. Wells, *Tibtech* 12: 173-184, 1994; Winter, G., *Drug Development Res.* 33: 71-89, 1994). In this example, a hG-CSF scanning permutagenesis library as described in Figure 1A is screened using the hG-CSF receptor expressed on mammalian cells as the affinity reagent. In Figure 1B, individual presented proteins are indicated by the shaded circles or diamonds and the affinity reagent is indicated by the light gray (pink) rectangles. Presentation

library particles are exposed to affinity reagent, unbound particles are washed away, and receptor-bound particles are eluted. The eluted particles are amplified in *E. coli*, and the screening cycle is repeated. During any round of the screening cycle, the genes encoded (in the present example encoding permuted proteins) by the selected particles can be expressed and evaluated.

Figure 2. Permuteins presented in the scanning permutagenesis library

Human G-CSF (ser17) protein is depicted as a string of circles, each circle corresponding to a single amino acid residue. Amino and carboxy ends of the protein are indicated. The amino acids of helical regions are indicated by medium gray balls, while the amino acids of inter-helical loops are indicated in light gray balls (See Hill et al., *Proc. Natl. Acad. Sci. USA* **90**: 5167-5171, 1993). Amino ends of the permuteins made for presentation in the library are indicated in dark gray. Asterisks indicate the breakpoints of the presented permuteins which were isolated by affinity screening with cells expressing hG-CSF receptor as illustrated in 1B.

Figure 3. Bioactivity of permuteins identified by affinity screening of the scanning permutagenesis library

Individual permuteins were expressed transiently in mammalian cells. Permeation molecules in the culture supernatants were quantitated by ELISA, and the proliferative activity of clones was determined using BAF-3-cells dependent on G-CSF for growth. The horizontal axis indicate concentration of protein and the vertical axis indicate incorporation of tritiated thymidine.

Definitions

The following is a list of abbreviations and the corresponding meanings as used interchangeably herein:

g = gram(s)
mg = milligram(s)
ml and mL = milliliter(s)
RT = room temperature
ug and µg = microgram(s)
uL and µl = microliter(s)

The following is a list of definitions of various terms used herein:

The term "permutein" means a circularly-permuted protein: a protein in which the amino and carboxy ends of the parent protein are joined together by a peptide linker sequence of zero or more amino acids. The amino and carboxy ends of the permuted protein occur at amino acids within the parental sequence.

The terms "chemical ligation" and "conjugation" mean a chemical reaction which covalently links two similar or dissimilar functional groups together intramolecularly or intermolecularly.

The term "peptide linker' means a compound which forms a carboxamide bond between two groups having one or more peptide linkages (CONH-) and serves as a connector for the propose of amelioration of the distance or space orientation between two molecules.

The term "native sequence" refers to an amino acid or nucleic acid sequence which is identical to a wild-type or native form of a gene or protein.

The terms "mutant amino acid sequence," "mutant protein", "variant protein", "mutein", or "mutant polypeptide" refer to a polypeptide having an amino acid sequence which varies from a native sequence due to amino acid additions, deletions, substitutions, or all three, or is encoded by a nucleotide sequence from an intentionally-made variant derived from a native sequence.

Detailed description of the invention

Determination of the amino and carboxyl termini of permuteins

The present invention encompasses circularly permuted-proteins of the formula C^1-L^1-N^1 prepared by phage display techniques. The polypeptide can be joined either directly or through a linker segment. The term "directly" defines permuteins in which the polypeptide ends are joined without a linker. Thus L^1 represents a chemical bond or a linker, preferably a polypeptide segment to which both C^1 and N^1 are joined, wherein C^1 is comprised of a segment derived from the carboxy portion of the parent protein and N^1 is comprised of a segment derived from the amino terminal portion of a parent protein represented by the general formula N^1-C^1. Preferably, N^1 and C^1 in the permuted protein C^1-L^1-N^1 are the same length as in the parent protein N^1-C^1, but each may be independently shorter or longer depending on the desired structural characteristics of the permutein. Most commonly L^1 is a linear peptide in which C^1 and N^1 are joined by amide

bonds, linking C^1 to the amino terminus of L^1 and carboxy terminus of L^1 to the amino terminus of N^1.

Additional peptide sequences may also be added to facilitate purification or identification of permuteins (e.g., poly-His). A highly antigenic peptide may also be added that would enable rapid assay and facile purification of the permuteins by a specific monoclonal antibody.

Determination of the linker

The linking group (L^1) is generally a polypeptide of between 1 and 500 amino acids in length. The linkers joining the two molecules are preferably designed to (1) allow the two molecules to fold and act independently of each other, (2) not have a propensity for developing an ordered secondary structure which could interfere with the functional domains of the two proteins, (3) have minimal hydrophobic characteristics which could interact with the functional protein domains and (4) provide steric separation of C^1 and N^1 such that C^1 and N^1 could interact simultaneously with their corresponding receptors on a single cell. Typically surface amino acids in flexible protein regions include Gly, Asn and Ser. Virtually any permutation of amino acid sequences containing Gly, Asn and Ser would be expected to satisfy the above criteria for a linker sequence. Other neutral amino acids, such as Thr and Ala, may also be used in the linker sequence. Additional amino acids may also be included in the linkers due to the addition of unique restriction sites in the linker sequence to facilitate construction of the multi-functional proteins.

Preferred L^1 linkers of the present invention include sequences selected from the group of formulas:

(SEQ ID NO: 1) through SEQ ID NO: 268)

Other linkers are also contemplated by the invention. The present invention is, however, not limited by the form, size or number of linker sequences employed. The only requirement of the linker is that it does not functionally interfere with the folding and function of the individual molecules of the multi-functional protein.

Utility of permuteins

Permuteins of the present invention may exhibit useful properties such as having similar or greater biological activity when compared to a single factor or by

having improved half-life or decreased adverse side effects, or a combination of these properties.

Permuteins which have little or no activity maybe useful as antigens for the production of antibodies for use in immunology or immunotherapy, as probes or as intermediates used to construct other useful permuteins.

The permuteins of the present invention may have an improved therapeutic profile as compared to their parent molecules. For example, some permuteins of the present invention may have a similar or more potent activities relative to other compounds or proteins without having a similar or corresponding increase in side-effects. This is particularly true of multifunctional or fusion protein therapeutics, where permutation may relieve steric and other hindrances that impair the activity of the parent fusion molecules (see Kreitman, R. J. et al., *Proc Natl Acad Sci USA* 91: 6889-6893, 1994; Kreitman, R. J. et al., *Cancer Res.* 55:3357-3363, 1995, for examples).

A general utility of permuteins is in the area of nanoscale devices described alternatively as "nanobiological" or "nanobiotechnological." These are nanoscale devices containing both precise structure nanomaterials and biological functional components (such as proteins). Nanodevices have been the subject of several reviews (Lee, S.C., *Trends in Biotechnology*, 16: 239-240, 1998).

Nanobiological/nanobiotechnological devices generally contain proteins covalently coupled to polymers or other non-biological precise structure materials. Issues of steric and other interferences with protein activity are applicable to proteins in nanobiological/nanobiotechnological devices and are highly analogous to the issues with multifunctional/fusion proteins discussed above. Protein permutation is fully expected to offer a viable approach to deal with these considerations, just as it does in the case of fusion proteins (Kreitman, R. J. et al., *Proc Natl Acad Sci USA* 91: 6889-6893, 1994; Kreitman, R. J. et al.,., *Cancer Res.* 55:3357-3363, 1995).

Examples

The following examples will illustrate the invention in greater detail, although it will be understood that the invention is not limited to these specific examples. Various other examples will be apparent to the person skilled in the art after reading the present disclosure without departing from the spirit and scope of

the invention. It is intended that all such other examples be included within the scope of the appended claims.

General Materials and Methods

General methods of cloning, expressing, and characterizing proteins are found in T. Maniatis et al., *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, 1982, and references cited therein, incorporated herein by reference; and in J. Sambrook et al., *Molecular Cloning, A Laboratory Manual*, 2nd edition, Cold Spring Harbor Laboratory, 1989, and references cited therein, incorporated herein by reference.

Unless noted otherwise, all specialty chemicals were obtained from Sigma Chemical Company (St. Louis, MO). Restriction endonucleases and T4 DNA ligase were obtained from New England Biolabs (Beverly, MA) or Boehringer Mannheim (Indianapolis, IN).

Strains, plasmids, and bacteriophage

The bacterial strains used in these studies are listed in Table 1. Plasmids and bacteriophage used or constructed in this study are listed in Tables 2 and 3, respectively.

Phage and phagemid stocks were made and manipulated as described (Kay, B.K., Winter, J., and McCafferty, J., Phage Display of Peptides and Proteins, Academic Press, San Diego, California, 1996; Merlin, S. et al., *Applied Biochemistry and Biotechnology* 67: 15-29, 1997).

Transformation of *E. coli* strains

E. coli strains (Table 1), such as DH5α™ (Life Technologies, Gaithersburg, MD) and TG1 (Amersham Corp., Arlington Heights, IL) are used for transformation of ligation reactions and are the hosts used to prepare plasmid DNA for transfecting mammalian cells. *E. coli* strains, such as JM101 (Yanisch-Perron et al., *Gene*, 33: 103-119, 1985) and MON105 (Obukowicz et al., *Appl. and Envir. Micr.*, 58: 1511-1523, 1992) can be used for expressing the multi-functional proteins of the present invention in the cytoplasm or periplasmic space.

DH5α™ Subcloning efficiency cells are purchased as competent cells and are ready for transformation using the manufacturer's protocol, while both *E. coli* strains TG1 and MON105 are rendered competent to take up DNA using a $CaCl_2$

method. Typically, 20 to 50 mL of cells are grown in LB medium (1% Bacto-tryptone, 0.5% Bacto-yeast extract, 150 mM NaCl) to a density of approximately 1.0 optical density unit at 600 nanometers (OD600) as measured by a Baush & Lomb Spectronic spectrophotometer (Rochester, NY). The cells are collected by centrifugation and resuspended in one-fifth culture volume of $CaCl_2$ solution (50 mM $CaCl_2$, 10 mM Tris-Cl, pH7.4) and are held at 4°C for 30 minutes. The cells are again collected by centrifugation and resuspended in one-tenth culture volume of $CaCl_2$ solution. Ligated DNA is added to 0.2 mL of these cells, and the samples are held at 4°C for 30-60 minutes. The samples are shifted to 42°C for two minutes and 1.0 mL of LB is added prior to shaking the samples at 37°C for one hour. Cells from these samples are spread on plates (LB medium plus 1.5% Bacto-agar) containing either ampicillin (100 micrograms/mL, ug/mL) when selecting for ampicillin-resistant transformants, or spectinomycin (75 ug/mL) when selecting for spectinomycin-resistant transformants. The plates are incubated overnight at 37°C.

Colonies are picked and inoculated into LB plus appropriate antibiotic (100 ug/mL ampicillin or 75 ug/mL spectinomycin) and are grown at 37°C while shaking.

DNA isolation and characterization

DNA constructs were made and propagated in *E. coli* using standard molecular biology techniques (Sambrook, J. et al., *Molecular Cloning, A Laboratory Manual,* 2nd edition, Cold Spring Harbor Laboratory, 1989).

Plasmid DNA can be isolated by a number of different methods and using commercially available kits known to those skilled in the art. Plasmid DNA is isolated using the Promega Wizard™ Miniprep kit (Madison, WI), the Qiagen QIAwell Plasmid isolation kits (Chatsworth, CA) or Qiagen Plasmid Midi or Mini kit. These kits follow the same general procedure for plasmid DNA isolation. Briefly, cells are pelleted by centrifugation (5000 x g), the plasmid DNA released with sequential NaOH/acid treatment, and cellular debris is removed by centrifugation (10000 x g). The supernatant (containing the plasmid DNA) is loaded onto a column containing a DNA-binding resin, the column is washed, and plasmid DNA eluted. After screening for the colonies with the plasmid of interest, the *E. coli* cells are inoculated into 50-100 ml of LB plus appropriate antibiotic for overnight growth at 37°C in an air incubator while shaking. The purified plasmid DNA is used for DNA sequencing, further restriction enzyme digestion, additional

subcloning of DNA fragments and transfection into *E. coli*, mammalian cells, or other cell types.

Sequence confirmation

DNA sequence analysis was performed using the Genesis 2000 DNA analysis system using standard methods (Prober et al., *Science* 238: 336-341, 1987).

Purified plasmid DNA is resuspended in dH_2O and its concentration is determined by measuring the absorbance at 260/280 nm in a Bausch and Lomb Spectronic 601 UV spectrometer. DNA samples are sequenced using ABI PRISM™ DyeDeoxy™ terminator sequencing chemistry (Applied Biosystems Division of Perkin Elmer Corporation, Lincoln City, CA) kits (Part Number 401388 or 402078) according to the manufacturer's suggested protocol usually modified by the addition of 5% DMSO to the sequencing mixture. Sequencing reactions are performed in a DNA thermal cycler (Perkin Elmer Corporation, Norwalk, CT) following the recommended amplification conditions. Samples are purified to remove excess dye terminators with Centri-Sep™ spin columns (Princeton Separations, Adelphia, NJ) and lyophilized. Fluorescent dye labeled sequencing reactions are resuspended in deionized formamide, and sequenced on denaturing 4.75% polyacrylamide-8M urea gels using ABI Model 373A and Model 377 automated DNA sequencers. Overlapping DNA sequence fragments are analyzed and assembled into master DNA contigs using Sequencher DNA analysis software (Gene Codes Corporation, Ann Arbor, MI).

Expression of permuted proteins in mammalian cells

To obtain sufficient protein for analysis of the bioactivity of individual MPO molecules containing cphG-CSFs, DNA segments containing individual affinity-selected MPO: cphGCSFs were subcloned into a mammalian expression vector, and expressed transiently in BHK cells as described below.

The BHK-21 cell line can be obtained from the ATCC (Rockville, MD). The cells are cultured in Dulbecco's modified Eagle media (DMEM/high-glucose), supplemented to 2 mM (mM) L-glutamine and 10% fetal bovine serum (FBS). This formulation is designated BHK growth media. Selective media is BHK growth media supplemented with 453 units/mL hygromycin B (CalBiochem, San Diego, CA). The BHK-21 cell line was previously stably transfected with the HSV transactivating protein VP16, which transactivates the IE110 promoter found on

the plasmid pMON3359 and pMON3633 and the IE175 promoter found in the plasmid pMON3360B (Hippenmeyer, P.J. and Pegg, L.E., *Curr. Opin. Biotechnol.* 6: 548-552, 1995). The VP16 protein drives expression of genes inserted behind the IE110 or IE175 promoter. BHK-21 cells expressing the transactivating protein VP16 are designated BHK-VP16. The plasmid pMON1118 expresses the hygromycin resistance gene from the SV40 promoter (Highkin et al., *Poultry Sci.*, 70: 970-981, 1991). A similar plasmid, pSV2-hph, is available from ATCC.

BHK-VP16 cells are seeded into a 60 millimeter (mm) tissue culture dish at 3 x 10^5 cells per dish 24 hours prior to transfection. Cells are transfected for 16 hours in 3 mL of "OPTIMEM"™ (Gibco-BRL, Gaithersburg, MD) containing 10 ug of plasmid DNA containing the gene of interest, 3 ug hygromycin resistance plasmid, pMON1118, and 80 ug of Gibco-BRL "LIPOFECTAMINE"™ per dish. The media is subsequently aspirated and replaced with 3 mL of growth media. At 48 hours post-transfection, media from each dish is collected and assayed for activity (transient conditioned media). The cells are removed from the dish by trypsin-EDTA, diluted 1:10, and transferred to 100 mm tissue culture dishes containing 10 mL of selective media. After approximately 7 days in selective media, resistant cells grow into colonies several millimeters in diameter. The colonies are removed from the dish with filter paper (cut to approximately the same size as the colonies and soaked in trypsin/EDTA) and transferred to individual wells of a 24 well plate containing 1 mL of selective media. After the clones are grown to confluence, the conditioned media is re-assayed, and positive clones are expanded into growth media.

Affinity selection and screening of phagemids

Affinity reagent used for the identification of functional MPO molecules containing cphG-CSF (MPO: cphG-CSF) species from the library were BHK cells expressing the hG-CSF receptor on their surface. The library pool was subjected to iterative affinity selection (four rounds) against BHK cells expressing the h-GCSF receptor using previously described techniques (Merlin, S. et al., *Applied Biochemistry and Biotechnology* 67: 15-29, 1997). Between rounds of selection, phage eluted from the affinity reagent were amplified in *E. coli* (Kay, B.K. J. Winter, and J. McCofferty, Phage Display of Peptides and Proteins, Academic Press, San Diego, California, 1996).

Expression of proteins in *E. coli*

When large-scale quantities of recombinant protein are desirable for structure-function studies, DNA segments containing individual affinity-selected MPO:cphGCSFs are subcloned into any of a variety of bacterial plasmid expression vectors, and expressed as a cytoplasmic product or as a secreted protein in *E. coli*.

E. coli strain MON105 or JM101 harboring the plasmid of interest are grown at 37°C in M9 plus casamino acids medium with shaking in an air incubator Model G25 from New Brunswick Scientific (Edison, NJ). Growth is monitored at OD_{600} until it reaches a value of 1.0 at which time nalidixic acid (10 mg/mL) in 0.1 N NaOH is added to a final concentration of 50 µg/mL, for cultures containing plasmids with the *E. coli recA* promoter driving expression of the recombinant gene. IPTG is used in place of nalidixic acid, as a chemical inducer to facilitate expression from plasmids containing the lac promoter or hybrid lac promoters. The cultures are then shaken at 37°C for three to four additional hours. A high degree of aeration is maintained throughout the culture period in order to achieve maximal production of the desired gene product. The cells are examined under a light microscope for the presence of inclusion bodies (IB). One mL aliquots of the culture are removed for analysis of protein content by boiling the pelleted cells, treating them with reducing buffer and electrophoresis via SDS-PAGE (see Maniatis et al., "Molecular Cloning: A Laboratory Manual", 1982). The culture is centrifuged (5000 x g) to pellet the cells.

Isolation of Inclusion Bodies

The cell pellet from a 330 mL *E. coli* culture is resuspended in 15 mL of sonication buffer (10 mM 2-amino-2-(hydroxymethyl) 1,3-propanediol hydrochloride (Tris-HCl), pH 8.0 + 1 mM ethylenediaminetetraacetic acid (EDTA). These resuspended cells are sonicated using the microtip probe of a Sonicator Cell Disruptor (Model W-375, Heat Systems-Ultrasonics, Inc., Farmingdale, New York). Three rounds of sonication in sonication buffer followed by centrifugation are employed to disrupt the cells and wash the inclusion bodies (IB). The first round of sonication is a 3 minute burst followed by a 1 minute burst, and the final two rounds of sonication are for 1 minute each.

Purification

The folded proteins can be affinity-purified using affinity reagents such as monoclonal antibodies or receptor subunits attached to a suitable matrix.

Purification can also be accomplished using any of a variety of chromatographic methods such as: ion exchange, gel filtration or hydrophobic chromatography or reversed phase HPLC. These and other protein purification methods are described in detail (Methods in Enzymology, Volume 182 "Guide to Protein Purification" edited by Murray Deutscher, Academic Press, San Diego, California, 1990).

Protein Characterization

The purified protein is analyzed by RP-HPLC, electrospray mass spectrometry, and SDS-PAGE. The protein quantitation is done by amino acid composition, RP-HPLC, and Bradford protein determination. In some cases tryptic peptide mapping is performed in conjunction with electrospray mass spectrometry to confirm the identity of the protein.

Baf-3/G-CSF receptor assay

Briefly, the mouse lymphoid cell line Baf3 was transfected with human granulocyte colony stimulating factor receptor (hG-CSFR) cDNA. Stable clones of Baf3 which expressed the G-CSFR and proliferated in the presence of hG-CSF were isolated and used to investigate the activity of human G-CSF receptor agonists without the influence of other human cytokine receptor responses.

The cDNA encoding hG-CSFR (a gift from Dr. Daniel C. Link (Washington University, St. Louis, MO) was released from the plasmid pEMCV.Sralpha as a *Hin*dIII/*Eco*RI (5' to 3') fragment, gel-purified, and inserted into the mammalian cell expression plasmid pcDNA3 (Invitrogen, San Diego, CA). This plasmid contains enhancer-promoter sequences from the immediate early gene of the human cytomegalovirus (CMV), a bovine growth hormone polyadenylation signal and transcription termination sequences, a neomycin resistance gene is present for the selection of G418 stable cell clones, and an ampicillin resistance gene for selection in *E. coli*. Ligation mixtures were transformed into *E. coli* strain TG1 [delta (*lac-pro*), *sup*E, *thi*, *hsd*D5/F'(*tra*D36, *pro*A$^+$B$^+$, lacI^q, *lac*ZdeltaM15] and plasmid DNA was purified using a Qiagen Midiprep Plasmid Kit. The structure of plasmid DNAs containing hG-CSFR were confirmed by restriction enzyme analysis and by automated DNA sequence analysis using an ABI sequencing machine. One of several plasmids with the correct structure was selected and given the designation pMON30298.

Baf3 cells, maintained in complete growth medium (RPMI 1640 supplemented to 10% FBS and 10% Wehi 3B supernatant as a source for mouse IL-

3), were seeded at a subconfluent cell density of 10^5 cells/ml in growth media (RPMI 1640 5% FBS; 2 mM L-glutamine) the day prior to the electroporation. The cells were collected and rinsed twice in 10 ml serum-free RPMI 1640. The cells were diluted to 10^6/ml in serum-free RPMI and 1 ml was placed into each electroporation chamber (Gibco/BRL #1608AJ). 50 ug of plasmid DNA was added to each chamber and the chambers were incubated on ice for 30 minutes prior to electroporation. The cells were electroporated on ice at a capacitance of 800 uF, 400V, fast charge, and low ohms in a BRL CellPorator. The cells were immediately removed from the chambers and placed into 10 cm dishes containing 10 ml of growth medium. The cells were allowed to recover for 48 hr in growth media prior to selection.

After the recovery period, the cells were pelleted at 1000 rpm for 10 minutes, and resuspended into 10 ml of selection medium (growth medium containing 800 ug/ml G418 sulfate (Gibco/BRL). The cells were kept in selection media, being passaged twice weekly, until only a few viable cells could be seen in the mock transfected control cell dishes (approximately 2 weeks). After an additional 2 weeks in selection media, the cells which had been electroporated with the hG-CSFR cDNA had grown to a cell density which allowed them to be tested for proliferation in the presence of hG-CSF (Fukunaga, R. et al., *EMBO J.* **10** (10): 2855-2865, 1991).

The cell proliferation assay conditions are as follows: Briefly, 25,000 cells were plated in a microtiter 96 well plate with or without cytokine in IMDM medium supplemented with BSA (50 ug/ml), human transferrin (100 ug/ml), lipid (50 ug/ml) 2-mercaptoethanol (50 uM final concentration). Each well was incubated with 0.5 uCi of ^{3}H-thymidine (16 hours) and the incorporated radioactivity was measured. Triplicate wells containing Baf3 cells were set up with 4 nM hG-CSF, 4 nM mIL-3 or media only control. Samples of different permuted proteins were tested in each assay.

Example 1: Construction of a permutein library without a linker region

Figure 1 shows a schematic of scanning permutagenesis. A plasmid construct comprising a tandem repeat of the modified human granulocyte colony stimulating factor (hG-CSF with a serine for amino acid 17) gene joined by a sequence (GCCGG, termed a zero order linker) was generated and subcloned into the plasmid pACYC177 (Chang, A.C.Y. and S.N. Cohen, *J Bacteriol.* **134**: 1141-1156, 1978) using standard molecular biology methods (Sambrook, J. et al.,

Molecular Cloning, A Laboratory Manual, 2nd edition, Cold Spring Harbor Press, New York, 1989). The resultant plasmid construct (pMON15978) was linearized by restriction digestion (*Sma*I) and used as a template for PCR amplification of circularly permuted hG-CSF (cphG-CSF) genes, following the method of Horlich (Horlick, R.A. et al., *Protein Engineering* 5: 427-431, 1992. For purposes of this demonstration of the scanning permutagenesis technique, we chose to make a limited permutein library rather than one containing every possible cphG-CSF. Figure 2 shows the position of the new amino termini for each new cphG-CSF.

Individual cphG-CSF genes were inserted into phagemid presentation vector pCANTAB 5E (Pharmacia Biotech,) such that they were expressed as a part of a MPO species (Feng, Y., N. R. Staten, C. M. Baum, N. L. Summers, M. Caparon, S. C. Bauer, L. Zurfluh, J. P. McKearn, B. K. Klein, S. C. Lee, C. A. McWherter. 1997. Multi-functional hematopoeitic receptor agonists. World Patent Application WO 97/12985) which was in turn fused to the amino end of the phage geneIII product. The presented fusion protein contained, starting from its amino terminus, a hIL-3 receptor agonist, cphG-CSF, and the phage gene III product. The juncture between the presented protein and the gene III product was as previously described (Merlin, S. et al., *Applied Biochemistry and Biotechnology* 67: 15-29, 1997).

After confirmation of the structure of each phagemid construct, phagemid particles were produced for each individual cphG-CSF-presenting species (Merlin et al., 1997). Some of these lots of particles were used to individually define the affinity properties of specific presented cphG-CSF species in analytical biopanning experiments (Caparon, M. H. et al., *Molecular Diversity* 1: 241-246, 1996; Merlin, S. et al., *Applied Biochemistry and Biotechnology* 67: 15-29, 1997), but all of the phage particle lots were titered and equivalent numbers of transducing units of each particle preparation were pooled together to form the scanning permutagenesis library for hG-CSF in an MPO background. Figure 2 shows the MPO: cp hG-CSF species present in the library.

MPO : cphG-CSF 38/37 is an example of the nomenclature used to specify the identity of individual permuted proteins. It describes a MPO molecule containing a circularly permuted human G-CSF module (with the serine 17 substitution). The first amino acid of the cphG-CSF domain is amino acid 38 of the parent protein, and the last amino acid is residue 37 of the parent.

Example 2: Presentation and Affinity screening of the MPO: cphGCSF

library

MPO: cphG-CSF 38/37, is a full hG-CSF receptor agonist (McKearn, J.P., Myelorestorative activities of synthokine and myelopoietin. *In* Proceedings of the 1996 IBC Conference on Therapeutic Applications of Cytokines, pp. 4.3.1-4.3.18, 1996). It was presented on filamentous phage as a positive control to demonstrate that permuted proteins can be presented on the surface of phage particles and affinity selected. After phagemid particles were produced from this construct, they were subjected to analytical biopanning using cells expressing the hG-CSF receptor as affinity reagent.

Table 1 shows that phage presented MPO: cphG-CSF 38/37 was affinity selected by cells expressing the hG-CSF. MPO: cphGCSF 38/37-GPIII fusion was expressed, secreted and assembled into phagemid particles, and could be affinity selected by the hG-CSF receptor. Permutagenesis of a protein does not appear to impair its successful presentation.

Relative to typical phage display libraries, the complexities of cp libraries are low, containing perhaps hundreds to thousands of individuals. The demonstration library here contained about 50 distinct clones, as opposed to more typical phage libraries containing more than 10^5 individuals (reviewed in Clackson, T. and J.A. Wells, *Tibtech* 12: 173-184, 1994).

37 randomly chosen selectants from round 1, and fewer from subsequent rounds (17, 11 and 14 were picked from rounds 2, 3 and 4, respectively) were chosen for sequence analysis. The identity of the MPO: cphG-CSFs identified in each round is shown in Figure 2.

A total of 14 MPO: cphGCSF species were identified from the output of affinity selection (Figure 2). Most of the MPO: cphGCSF species identified from the library had new carboxy and amino termini in loop segments (9 of 14 permuteins identified), rather than in clearly defined secondary structures (See Hill et al., 1993 for the hG-CSF structure). Five selectants had termini within helical domains of hG-CSF (MPO: cphG-CSFs 13/12, 19/18, 71/70, 123/122 and 159/158). For three of these molecules (MPO: cphG-CSFs 13/12, 71/70 and 123/122) their new ends lie at the outermost ends of helices, and therefore perturbation of secondary structure caused by these permuteins may be minimal. However, MPO: cphG-CSF 19/18 and MPO: cphG-CSF 159/158 have new termini well within helix 1 and helix 4 of hG-CSF, respectively.

These data parallel the observations of Graf and Schachman, who developed a limited DNase I digestion method for "random" permutagenesis (Graf, R. and H. K. Schachman, *Proc Natl Acad Sci USA* 93:11591-11596, 1996). They identified two permutein species of aspartate transcarbamoylase that introduced new amino and carboxy ends into secondary structural domains and that retained biological activity. In their work, the majority of permuteins introducing ends into secondary structures (5/7 identified) were significantly diminished in activity. In contrast, we found a several permuteins that introduced helical breaks retained activity (See Below). The method used by Graf and Schechman frequently introduces point mutations, small insertions and deletions into the permuted proteins, potentially complicating the analysis of the effects of permutagenesis.

Example 3: Biological activity of MPO: cphG-CSFs selected from the cp phage library

To obtain sufficient protein for analysis of the bioactivity of individual MPO molecules containing cphG-CSFs, DNA segments containing individual affinity-selected MPO: cphGCSFs were subcloned into a mammalian expression vector, and expressed transiently in BHK cells as described above.

The MPO: cphG-CSFs isolated from biopanning were all expressed transiently in mammalian cells and the amount of MPO: cphG-CSF in each supernatant was determined by sandwich hIl-3 ELISA (Olins P.O. et al., *J. Biol. Chem.* 270: 23754-23760, 1995). The quantitated supernatants were then assayed for G-CSF receptor agonist activity in a Baf-3/G-CSF receptor assay (Figure 3, Table 4).

All but one of the transiently expressed MPO: cphG-CSF proteins exhibited G-CSF activity equivalent to or slightly better than that of the parent MPO molecule, including those MPO: cphG-CSFs with new carboxy and amino ends within helixes. The permutein encoded by pMON16021 with a breakpoint between positions 48 and 49 did not exhibit activity in the G-CSF-dependent proliferation assay. These data suggest that most of the proteins isolated from the library are competent to bind the hG-CSF receptor and produce a proliferation signal.

All references, patents, or applications cited herein are incorporated by reference in their entirety, as if written herein.

References

Buchwalder, A. et al., *Biochemistry* 31:1621-1630, 1992.

Caparon, M. H. et al., *Molecular Diversity* 1: 241-246, 1996.

Chang, A.C.Y. and S.N. Cohen, *J Bacteriol.* 134: 1141-1156, 1978.

Clackson, T. and J.A. Wells, *Tibtech* 12: 173-184, 1994.

Feng et al., *J. Mol. Biol.* 259: 524-551, 1996.

Feng, Y., N. R. Staten, C. M. Baum, N. L. Summers, M. Caparon, S. C. Bauer, L. Zurfluh, J. P. McKearn, B. K. Klein, S. C. Lee, C. A. McWherter. 1997. Multi-functional hematopoeitic receptor agonists. World Patent Application WO 97/12985.

Fukunaga, R., Ishizaka-Ikeda, E., Pan, C. X., Seto, Y., and Nagata, S. Functional Domains of the Granulocyte Colony-Stimulating Factor Receptor. *EMBO J.* 10 (10):2855-2865, 1991.

Goldenberg, D. P. and T. E. Creighton, *J. Mol. Biol.* 165: 407-413, 1983.

Graf, R. and H. K. Schachman, *Proc Natl Acad Sci USA* 93:11591-11596, 1996.

Hahn, M. et al., *Proc Natl Acad Sci USA* 91: 10417-10421, 1994.

Highkin et al., *Poultry Sci.*, 70: 970-981, 1991.

Hill et al., *Proc. Natl. Acad. Sci. USA* 90: 5167-5171, 1993.

Hippenmeyer, P.J. and L.E. Pegg, *Curr. Opin. Biotechnol.* 6: 548-552, 1995.

Horlick, R.A. et al., *Protein Engineering* 5: 427-431, 1992.

Jelinski, L.W. Biologically related aspects of nanoparticles, nanostructured materials and nanodevices. *In* "WTEC workshop on Global Assessment of R &D Status and Trends in Nanoparticles, Nanostructured Materials and Nanodevices", International Technology Research Institute, Loyola College, Baltimore, MD., S. C., 1998.

Johnson, J. and F. M. Raushel, *Biochemistry* 35: 10223-10233, 1996.

Kay, B.K. J. Winter, and J. McCofferty, Phage Display of Peptides and Proteins, Academic Press, San Diego, California, 1996.

Koebnik, R. and L. Kramer, *J. Mol. Biol.* **250**: 617-626, 1995.

Kreitman, R. J. et al. *Cytokine* **7**(4): 311-318, 1995.

Kreitman, R. J. et al., *Proc Natl Acad Sci USA* **91**: 6889-6893, 1994.

Kreitman, R. J. et al., *Cancer Res.* **55**:3357-3363, 1995.

Lee, S.C. Biotechnology for Nanotechnology. *Trends in Biotechnology*, **16**: 239-240, 1998.

Lee, S.C., Phage presentation for cytokine engineering. IBC's Second International Conference on Display Technologies, 1997.

Lin, X. et al., *Protein Science* **4**: 159-166, 1995.

Lowman and Wells, *J. Mol. Biol.* **234**: 564-578, 1993.

Luger et al., *Protein Engineering* **3**: 249-258, 1990.

Luger et al., *Science* **243**: 206-210, 1989.

Maniatis, T. et al., *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, 1982.

McKearn, J.P. Myelorestorative activities of synthokine and myelopoietin. In Proceedings of the 1996 IBC Conference on Therapeutic Applications of Cytokines, pp. 4.3.1-4.3.18, 1996.

Merlin, S. et al., *Applied Biochemistry and Biotechnology* **67**: 15-29, 1997.

Mullins, L. S. et al., *J Am. Chem Soc.* **116**: 5529-5533, 1994.

Murray Deutscher (ed), Methods in Enzymology, Volume 182 "Guide to Protein Purification," Academic Press, San Diego, California, 1990.

Obukowicz, M. et al., *Appl. and Envir. Micr.*, 58: 1511-1523, 1992.

Olins P.O. et al., *J. Biol. Chem.* **270**: 23754-23760, 1995.

Prober, J.M. et al., *Science* **238**: 336-341, 1987.

Protosova, N. Y. et al., *Protein Engineering* 7: 1373-1377, 1994.

Puri, R. K. et al., *Cellular Immunol.* 171: 80-86, 1996.

Rose, K. et al., *Molecular Immunology* 32: 1031-1037, 1995.

Sambrook, J. et al., *Molecular Cloning, A Laboratory Manual,* 2nd edition, Cold Spring Harbor Press, New York, 1989.

Shortle, D. and J. Sondek, *Curr. Opin. Biotechnol.* 6: 299-305.

Smith, G. P. , *Curr. Opin. in Biotechnol.* 2: 668-673, 1991.

Smith, G. P., *Science* 228: 1315-1317, 1985.

Sondek, J. and D. Shortle, *Proteins* 7: 387-393, 1990.

Thomas, J. W. et al., *Proc Natl Acad Sci USA* 92: 3779-3783, 1995.

Winter, G., *Drug Development Res.* 33: 71-89, 1994.

Yang, Y. R. and H. K. Schachman, *Proc Natl Acad Sci USA* 90: 11980-11984, 1993.

Yanisch-Perron et al., *Gene*, 33: 103-119, 1985.

Zhang, T. et al., *Biochemistry* 32: 12311-12318, 1993.

Tables

Table 1: Circularly permuted proteins

Protein	Reference
Enzymes	
T4 lysozyme	Zhang et al., *Biochemistry* **32**:12311-12318 (1993); Zhang et al., *Nature Struct. Biol.* **1**:434-438 (1995)
dihydrofolate reductase	Buchwalder et al., *Biochemistry* **31**:1621-1630 (1994); Protasova et al., *Prot. Eng.* **7**:1373-1377 (1995)
ribonuclease T1	Mullins et al., *J. Am. Chem. Soc.* **116**:5529-5533 (1994); Garrett et al., *Protein Science* **5**:204-211 (1996)
Bacillus β-glucanase	Hahn et al., *Proc. Natl. Acad. Sci. U.S.A.* **91**:10417-10421 (1994)
aspartate transcarbamoylase	Yang and Schachman, *Proc. Natl. Acad.* Transcarbamoylase *Sci. U.S.A.* **90**:11980-11984 (1993)
phosphoribosyl anthranilate isomerase	Luger et al., *Science* **243**:206-210 (1989); Luger et al., *Prot. Eng.* **3**:249-258 (1990)
pepsin/pepsinogen	Lin et al., *Protein Science* **4**:159-166 (1995)
glyceraldehyde-3-phosphate dehydrogenase	Vignais et al., *Protein Science* **4**:994-1000 (1995)
ornithine decarboxylase	Li & Coffino, *Mol. Cell. Biol.* **13**:2377-2383 (1993)
yeast phosphoglycerate dehydrogenase	Ritco-Vonsovici et al., *Biochemistry* **34**:16543-16551 (1995)
Enzyme Inhibitor	
basic pancreatic trypsin inhibitor	Goldenberg & Creighton, *J. Mol. Biol.* **165**:407-413 (1983)
Cytokines	
interleukin-1β	Horlick et al., *Protein Eng.* **5**:427-431 (1992)
interleukin-4	Kreitman et al., *Cytokine* **7**:311-318 (1995)
Tyrosine Kinase Recognition Domain	

α-spectrin SH3 domain	Viguera et al., *J. Mol. Biol.* 247:670-681 (1995)
Transmembrane Protein	
omp A	Koebnik & Krämer, *J. Mol. Biol.* 250:617-626 (1995)
Chimeric Protein	
interleukin-4-*Pseudomonas* exotoxin fusion molecule	Kreitman et al., *Proc. Natl. Acad. Sci. U.S.A.* 91:6889-6893 (1994)

Table 2: Strains

Designation	Description or Genotype	Reference/Source
DH5α™	F⁻, *phi*80 d*lac*ZdeltaM15, delta(*lac*ZYA-*arg*F)U169, *deo*R, *rec*A1, *end*A1, *hsd*R17 (rk⁻,mk⁺), *pho*A, *sup*E44, lambda-, *thi*-1, *gyr*A96, relA1	Life Technologies, Rockville, Maryland
JM101 (ATCC# 33876)	delta (*pro lac*), *sup*E, *thi*, F'(*tra*D36, *pro*A⁺B⁺, *lac*I^q, *lac*ZdeltaM15)	Yanisch-Perron et al., *Gene*, 33: 103-119, 1985
MON105 (ATCC# 55204)	F⁻, lambda-,IN (rrnD, rrnE)1, rpoD⁺, *rpo*H358	Obukowicz et al., *Appl. and Envir. Micr.*, 58: 1511-1523, 1992
MON208	W3110 *rpo*H358, *lac*I^q, *omp*T::*kan*	Alan Easton
TG1	delta(lac-pro), *sup*E, *thi*-1, *hsd*D5/F'(*tra*D36, *pro*A⁺B⁺, *lac*I^q, *lac*ZdeltaM15)	Amersham Corp., Arlington Heights, Illinois
W3110	IN (*rrn*D-*rrn*E)1, *rph*1	Lab collection

Table 3: Plasmids

Plasmid	SEQ ID NO.	Selectable Marker	Description	Source
pACYC177		Kan^R Amp^R	Plasmid with multiple cloning sites and two selectable markers	Chang, A.C.Y. and S.N. Cohen, *J Bacteriol.* 134: 1141-1156, 1978
pMON15978		Amp^R	Plasmid construct comprising a tandem repeat of the modified human granulocyte colony stimulating factor (hG-CSF with a serine for amino acid 17) gene joined by a sequence (GCCGG, termed a zero order linker), subcloned into the plasmid pACYC177	This work
pCANTAB 5E		Amp^R	Phage display vector containing lac promoter operably linked to fd gene 3 signal sequence, a linker region, an E tag, and an fd gene 3 structural gene all cloned into the vector backbone of pUC119 containing ColE1 ori, the beta lactamase resistance gene, and an M13 ori.	Pharmacia Biotech, Piscataway, NJ
pMON16016		Amp^R	Phagemid presentation vector pCANTAB 5E derivation containing inserted individual cphG-CSF gene such that it was expressed as a part of an	This work

		MPO species, fused in turn to the amino terminus end of the phage geneIII product. The first amino acid of the cphG-CSF domain is amino acid 1 of the parent, and the last amino acid is residue 174 of the parent. The zero order linker is attached at the carboxyl end of amino acid 174.	
pMON16017	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 3 of the parent, and the last amino acid is residue 2 of the parent.	This work
pMON16029	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 7 of the parent, and the last amino acid is residue 6 of the parent.	This work
pMON16030	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 9 of the parent, and the last amino acid is residue 8 of the parent.	This work
pMON16018	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 11 of the parent, and the last amino acid is residue 10 of the parent.	This work

pMON16019	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 13 of the parent, and the last amino acid is residue 12 of the parent.	This work
pMON16031	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 15 of the parent, and the last amino acid is residue 14 of the parent.	This work
pMON16020	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 19 of the parent, and the last amino acid is residue 18 of the parent.	This work
pMON16032	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 22 of the parent, and the last amino acid is residue 21 of the parent.	This work
pMON16033	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 27 of the parent, and the last amino acid is residue 26 of the parent.	This work
pMON16034	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 31 of the parent, and the	This work

		last amino acid is residue 30 of the parent.	
pMON16035	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 35 of the parent, and the last amino acid is residue 34 of the parent.	This work
pMON16036	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 37 of the parent, and the last amino acid is residue 36 of the parent.	This work
pMON16037	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 38 of the parent, and the last amino acid is residue 37 of the parent.	This work
pMON16038	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 39 of the parent, and the last amino acid is residue 38 of the parent.	This work
pMON16039	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 43 of the parent, and the last amino acid is residue 42 of the parent.	This work
pMON16040	Amp^R	Identical to pMON16016 except the first amino acid of the	This work

		cphG-CSF domain is amino acid 45 of the parent, and the last amino acid is residue 44 of the parent.	
pMON16041	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 47 of the parent, and the last amino acid is residue 46 of the parent.	This work
pMON16022	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 49 of the parent, and the last amino acid is residue 48 of the parent.	This work
pMON16042	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 51 of the parent, and the last amino acid is residue 50 of the parent.	This work
pMON16043	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 53 of the parent, and the last amino acid is residue 52 of the parent.	This work
pMON16044	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 56 of the parent, and the last amino acid is residue 55 of the parent.	This work

pMON16023	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 60 of the parent, and the last amino acid is residue 59 of the parent.	This work
pMON16045	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 64 of the parent, and the last amino acid is residue 63 of the parent.	This work
pMON16024	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 67 of the parent, and the last amino acid is residue 66 of the parent.	This work
pMON16046	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 69 of the parent, and the last amino acid is residue 68 of the parent.	This work
pMON16025	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 71 of the parent, and the last amino acid is residue 70 of the parent.	This work
pMON16047	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 73 of the parent, and the	This work

		last amino acid is residue 72 of the parent.	
pMON16048	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 84 of the parent, and the last amino acid is residue 83 of the parent.	This work
pMON16049	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 98 of the parent, and the last amino acid is residue 97 of the parent.	This work
pMON16050	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 100 of the parent, and the last amino acid is residue 99 of the parent.	This work
pMON16051	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 102 of the parent, and the last amino acid is residue 101 of the parent.	This work
pMON16052	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 112 of the parent, and the last amino acid is residue 111 of the parent.	This work
pMON16053	Amp^R	Identical to pMON16016 except the first amino acid of the	This work

		cphG-CSF domain is amino acid 121 of the parent, and the last amino acid is residue 120 of the parent.	
pMON16026	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 123 of the parent, and the last amino acid is residue 122 of the parent.	This work
pMON16027	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 125 of the parent, and the last amino acid is residue 124 of the parent.	This work
pMON16054	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 133 of the parent, and the last amino acid is residue 132 of the parent.	This work
pMON16055	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 142 of the parent, and the last amino acid is residue 141 of the parent.	This work
pMON16056	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 143 of the parent, and the last amino acid is residue 142 of the parent.	This work

pMON16057	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 147 of the parent, and the last amino acid is residue 146 of the parent.	This work
pMON16028	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 159 of the parent, and the last amino acid is residue 158 of the parent.	This work
pMON16058	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 168 of the parent, and the last amino acid is residue 167 of the parent.	This work
pMON16059	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 170 of the parent, and the last amino acid is residue 169 of the parent.	This work

Table 4: Analytical biopanning

Before receptor*	After receptor*	Enrichment
$1/6.6x10^4$	$1/6.5x10^{-1}$	990-fold

* Amp^R/Kan^R resistant colonies

Analytical biopanning shows that MPO molecules containing permuted hG-CSF domains can be presented and affinity selected in a hG-CSF receptor dependent fashion. A mixture of phagemids presenting MPO: cphG-CSF 38/37 (ampicillin resistant) and M13k07 (kanamycin resistant) were exposed to BHK cells with or without the hG-CSF receptor on their surface, washed and eluted from the cell surface. Eluted phage were introduced into *E. coli* and the transfected cells were plated on media containing kanamycin or ampicillin. The ratio of ampicillin resistant to kanamycin resistant particles were determined prior to and following exposure to receptor by counting resistant colonies.

Table 5: Activity of selected permuteins

Plasmid	Permutein breakpoint in G-CSF amino acid sequence	Activity in G-CSF-dependent proliferation assay
pMON16017	3/2	+
pMON16018	11/10	+
pMON16019	13/12	+
pMON16020	19/18	+
pMON16021	49/48	–
pMON16022	60/59	+
pMON16023	67/66	+
pMON16024	69/68	+
pMON16025	71/70	+
pMON16026	123/122	+
pMON16027	125/124	+
pMON16028	159/158	+

Table 6: SEQ ID Number/SEQ ID Name Correlation

SEQ ID NO.	SEQ ID Name	Sequence									
1.	FGS1	CGCGCGC	ACATG	TCT	ACA	CCA	TTG	GGC	CCT	GCC	AGC
2.	FGS2	CGCGCGC	ACATG	TCT	CCA	TTG	GGC	CCT	GCC	AGC	TCC
3.	FGS3	CGCGCGC	ACATG	TCT	TTG	GGC	CCT	GCC	AGC	TCC	CTG
4.	FGS4	CGCGCGC	ACATG	TCT	GGC	CCT	GCC	AGC	TCC	CTG	CCC
5.	FGS5	CGCGCGC	ACATG	TCT	CCT	GCC	AGC	TCC	CTG	CCC	CAG
6.	FGS6	CGCGCGC	ACATG	TCT	GCC	AGC	TCC	CTG	CCC	CAG	AGC
7.	FGS7	CGCGCGC	ACATG	TCT	AGC	TCC	CTG	CCC	CAG	AGC	TTC
8.	FGS8	CGCGCGC	ACATG	TCT	TCC	CTG	CCC	CAG	AGC	TTC	CTG
9.	FGS9	CGCGCGC	ACATG	TCT	CTG	CCC	CAG	AGC	TTC	CTG	CTC
10.	FGS10	CGCGCGC	ACATG	TCT	CCC	CAG	AGC	TTC	CTG	CTC	AAG
11.	FGS11	CGCGCGC	ACATG	TCT	CAG	AGC	TTC	CTG	CTC	AAG	TCT
12.	FGS12	CGCGCGC	ACATG	TCT	AGC	TTC	CTG	CTC	AAG	TCT	TTA
13.	FGS13	CGCGCGC	ACATG	TCT	TTC	CTG	CTC	AAG	TCT	TTA	GAG
14.	FGS14	CGCGCGC	ACATG	TCT	CTG	CTC	AAG	TCT	TTA	GAG	CAA
15.	FGS15	CGCGCGC	ACATG	TCT	CTC	AAG	TCT	TTA	GAG	CAA	GTG
16.	FGS16	CGCGCGC	ACATG	TCT	AAG	TCT	TTA	GAG	CAA	GTG	AGG
17.	FGS17	CGCGCGC	ACATG	TCT	TCT	TTA	GAG	CAA	GTG	AGG	AAG
18.	FGS18	CGCGCGC	ACATG	TCT	TTA	GAG	CAA	GTG	AGG	AAG	ATC
19.	FGS19	CGCGCGC	ACATG	TCT	GAG	CAA	GTG	AGG	AAG	ATC	CAG
20.	FGS20	CGCGCGC	ACATG	TCT	CAA	GTG	AGG	AAG	ATC	CAG	GGC
21.	FGS21	CGCGCGC	ACATG	TCT	GTG	AGG	AAG	ATC	CAG	GGC	GAT
22.	FGS22	CGCGCGC	ACATG	TCT	AGG	AAG	ATC	CAG	GGC	GAT	GGC
23.	FGS23	CGCGCGC	ACATG	TCT	AAG	ATC	CAG	GGC	GAT	GGC	GCA
24.	FGS24	CGCGCGC	ACATG	TCT	ATC	CAG	GGC	GAT	GGC	GCA	GCG
25.	FGS25	CGCGCGC	ACATG	TCT	CAG	GGC	GAT	GGC	GCA	GCG	CTC
26.	FGS26	CGCGCGC	ACATG	TCT	GGC	GAT	GGC	GCA	GCG	CTC	CAG
27.	FGS27	CGCGCGC	ACATG	TCT	GAT	GGC	GCA	GCG	CTC	CAG	GAG
28.	FGS28	CGCGCGC	ACATG	TCT	GGC	GCA	GCG	CTC	CAG	GAG	AAG
29.	FGS29	CGCGCGC	ACATG	TCT	GCA	GCG	CTC	CAG	GAG	AAG	CTG
30.	FGS30	CGCGCGC	ACATG	TCT	GCG	CTC	CAG	GAG	AAG	CTG	TGT
31.	FGS31	CGCGCGC	ACATG	TCT	CTC	CAG	GAG	AAG	CTG	TGT	GCC
32.	FGS32	CGCGCGC	ACATG	TCT	CAG	GAG	AAG	CTG	TGT	GCC	ACC
33.	FGS33	CGCGCGC	ACATG	TCT	GAG	AAG	CTG	TGT	GCC	ACC	TAC
34.	FGS34	CGCGCGC	ACATG	TCT	AAG	CTG	TGT	GCC	ACC	TAC	AAG
35.	FGS35	CGCGCGC	ACATG	TCT	CTG	TGT	GCC	ACC	TAC	AAG	CTG
36.	FGS36	CGCGCGC	ACATG	TCT	TGT	GCC	ACC	TAC	AAG	CTG	TGC
37.	FGS37	CGCGCGC	ACATG	TCT	GCC	ACC	TAC	AAG	CTG	TGC	CAC
38.	FGS38	CGCGCGC	ACATG	TCT	ACC	TAC	AAG	CTG	TGC	CAC	CCC
39.	FGS39	CGCGCGC	ACATG	TCT	TAC	AAG	CTG	TGC	CAC	CCC	GAG
40.	FGS40	CGCGCGC	ACATG	TCT	AAG	CTG	TGC	CAC	CCC	GAG	GAG
41.	FGS41	CGCGCGC	ACATG	TCT	CTG	TGC	CAC	CCC	GAG	GAG	CTG
42.	FGS42	CGCGCGC	ACATG	TCT	TGC	CAC	CCC	GAG	GAG	CTG	GTG
43.	FGS43	CGCGCGC	ACATG	TCT	CAC	CCC	GAG	GAG	CTG	GTG	CTG
44.	FGS44	CGCGCGC	ACATG	TCT	CCC	GAG	GAG	CTG	GTG	CTG	CTC
45.	FGS45	CGCGCGC	ACATG	TCT	GAG	GAG	CTG	GTG	CTG	CTC	GGA
46.	FGS46	CGCGCGC	ACATG	TCT	GAG	CTG	GTG	CTG	CTC	GGA	CAC
47.	FGS47	CGCGCGC	ACATG	TCT	CTG	GTG	CTG	CTC	GGA	CAC	TCT
48.	FGS48	CGCGCGC	ACATG	TCT	GTG	CTG	CTC	GGA	CAC	TCT	CTG
49.	FGS49	CGCGCGC	ACATG	TCT	CTG	CTC	GGA	CAC	TCT	CTG	GGC
50.	FGS50	CGCGCGC	ACATG	TCT	CTC	GGA	CAC	TCT	CTG	GGC	ATC
51.	FGS51	CGCGCGC	ACATG	TCT	GGA	CAC	TCT	CTG	GGC	ATC	CCC
52.	FGS52	CGCGCGC	ACATG	TCT	CAC	TCT	CTG	GGC	ATC	CCC	TGG
53.	FGS53	CGCGCGC	ACATG	TCT	TCT	CTG	GGC	ATC	CCC	TGG	GCT
54.	FGS54	CGCGCGC	ACATG	TCT	CTG	GGC	ATC	CCC	TGG	GCT	CCC
55.	FGS55	CGCGCGC	ACATG	TCT	GGC	ATC	CCC	TGG	GCT	CCC	CTG
56.	FGS56	CGCGCGC	ACATG	TCT	ATC	CCC	TGG	GCT	CCC	CTG	AGC
57.	FGS57	CGCGCGC	ACATG	TCT	CCC	TGG	GCT	CCC	CTG	AGC	TCC
58.	FGS58	CGCGCGC	ACATG	TCT	TGG	GCT	CCC	CTG	AGC	TCC	TGC
59.	FGS59	CGCGCGC	ACATG	TCT	GCT	CCC	CTG	AGC	TCC	TGC	CCC
60.	FGS60	CGCGCGC	ACATG	TCT	CCC	CTG	AGC	TCC	TGC	CCC	AGC
61.	FGS61	CGCGCGC	ACATG	TCT	CTG	AGC	TCC	TGC	CCC	AGC	CAG
62.	FGS62	CGCGCGC	ACATG	TCT	AGC	TCC	TGC	CCC	AGC	CAG	GCC
63.	FGS63	CGCGCGC	ACATG	TCT	TCC	TGC	CCC	AGC	CAG	GCC	CTG
64.	FGS64	CGCGCGC	ACATG	TCT	TGC	CCC	AGC	CAG	GCC	CTG	CAG
65.	FGS65	CGCGCGC	ACATG	TCT	CCC	AGC	CAG	GCC	CTG	CAG	CTG
66.	FGS66	CGCGCGC	ACATG	TCT	AGC	CAG	GCC	CTG	CAG	CTG	GCA
67.	FGS67	CGCGCGC	ACATG	TCT	CAG	GCC	CTG	CAG	CTG	GCA	GGC
68.	FGS68	CGCGCGC	ACATG	TCT	GCC	CTG	CAG	CTG	GCA	GGC	TGC
69.	FGS69	CGCGCGC	ACATG	TCT	CTG	CAG	CTG	GCA	GGC	TGC	TTG
70.	FGS70	CGCGCGC	ACATG	TCT	CAG	CTG	GCA	GGC	TGC	TTG	AGC
71.	FGS71	CGCGCGC	ACATG	TCT	CTG	GCA	GGC	TGC	TTG	AGC	CAA
72.	FGS72	CGCGCGC	ACATG	TCT	GCA	GGC	TGC	TTG	AGC	CAA	CTC
73.	FGS73	CGCGCGC	ACATG	TCT	GGC	TGC	TTG	AGC	CAA	CTC	CAT
74.	FGS74	CGCGCGC	ACATG	TCT	TGC	TTG	AGC	CAA	CTC	CAT	AGC
75.	FGS75	CGCGCGC	ACATG	TCT	TTG	AGC	CAA	CTC	CAT	AGC	GGC
76.	FGS76	CGCGCGC	ACATG	TCT	AGC	CAA	CTC	CAT	AGC	GGC	CTT
77.	FGS77	CGCGCGC	ACATG	TCT	CAA	CTC	CAT	AGC	GGC	CTT	TTC
78.	FGS78	CGCGCGC	ACATG	TCT	CTC	CAT	AGC	GGC	CTT	TTC	CTC
79.	FGS79	CGCGCGC	ACATG	TCT	CAT	AGC	GGC	CTT	TTC	CTC	TAC

80.	FGS80	CGCGCGC	ACATG	TCT	AGC	GGC	CTT	TTC	CTC	TAC	CAG
81.	FGS81	CGCGCGC	ACATG	TCT	GGC	CTT	TTC	CTC	TAC	CAG	GGG
82.	FGS82	CGCGCGC	ACATG	TCT	CTT	TTC	CTC	TAC	CAG	GGG	CTC
83.	FGS83	CGCGCGC	ACATG	TCT	TTC	CTC	TAC	CAG	GGG	CTC	CTG
84.	FGS84	CGCGCGC	ACATG	TCT	CTC	TAC	CAG	GGG	CTC	CTG	CAG
85.	FGS85	CGCGCGC	ACATG	TCT	TAC	CAG	GGG	CTC	CTG	CAG	GCC
86.	FGS86	CGCGCGC	ACATG	TCT	CAG	GGG	CTC	CTG	CAG	GCC	CTG
87.	FGS87	CGCGCGC	ACATG	TCT	GGG	CTC	CTG	CAG	GCC	CTG	GAA
88.	FGS88	CGCGCGC	ACATG	TCT	CTC	CTG	CAG	GCC	CTG	GAA	GGG
89.	FGS89	CGCGCGC	ACATG	TCT	CTG	CAG	GCC	CTG	GAA	GGG	ATA
90.	FGS90	CGCGCGC	ACATG	TCT	CAG	GCC	CTG	GAA	GGG	ATA	TCC
91.	FGS91	CGCGCGC	ACATG	TCT	GCC	CTG	GAA	GGG	ATA	TCC	CCC
92.	FGS92	CGCGCGC	ACATG	TCT	CTG	GAA	GGG	ATA	TCC	CCC	GAG
93.	FGS93	CGCGCGC	ACATG	TCT	GAA	GGG	ATA	TCC	CCC	GAG	TTG
94.	FGS94	CGCGCGC	ACATG	TCT	GGG	ATA	TCC	CCC	GAG	TTG	GGT
95.	FGS95	CGCGCGC	ACATG	TCT	ATA	TCC	CCC	GAG	TTG	GGT	CCC
96.	FGS96	CGCGCGC	ACATG	TCT	TCC	CCC	GAG	TTG	GGT	CCC	ACC
97.	FGS97	CGCGCGC	ACATG	TCT	CCC	GAG	TTG	GGT	CCC	ACC	TTG
98.	FGS98	CGCGCGC	ACATG	TCT	GAG	TTG	GGT	CCC	ACC	TTG	GAC
99.	FGS99	CGCGCGC	ACATG	TCT	TTG	GGT	CCC	ACC	TTG	GAC	ACA
100.	FGS100	CGCGCGC	ACATG	TCT	GGT	CCC	ACC	TTG	GAC	ACA	CTG
101.	FGS101	CGCGCGC	ACATG	TCT	CCC	ACC	TTG	GAC	ACA	CTG	CAG
102.	FGS102	CGCGCGC	ACATG	TCT	ACC	TTG	GAC	ACA	CTG	CAG	CTG
103.	FGS103	CGCGCGC	ACATG	TCT	TTG	GAC	ACA	CTG	CAG	CTG	GAC
104.	FGS104	CGCGCGC	ACATG	TCT	GAC	ACA	CTG	CAG	CTG	GAC	GTC
105.	FGS105	CGCGCGC	ACATG	TCT	ACA	CTG	CAG	CTG	GAC	GTC	GCC
106.	FGS106	CGCGCGC	ACATG	TCT	CTG	CAG	CTG	GAC	GTC	GCC	GAC
107.	FGS107	CGCGCGC	ACATG	TCT	CAG	CTG	GAC	GTC	GCC	GAC	TTT
108.	FGS108	CGCGCGC	ACATG	TCT	CTG	GAC	GTC	GCC	GAC	TTT	GCC
109.	FGS109	CGCGCGC	ACATG	TCT	GAC	GTC	GCC	GAC	TTT	GCC	ACC
110.	FGS110	CGCGCGC	ACATG	TCT	GTC	GCC	GAC	TTT	GCC	ACC	ACC
111.	FGS111	CGCGCGC	ACATG	TCT	GCC	GAC	TTT	GCC	ACC	ACC	ATC
112.	FGS112	CGCGCGC	ACATG	TCT	GAC	TTT	GCC	ACC	ACC	ATC	TGG
113.	FGS113	CGCGCGC	ACATG	TCT	TTT	GCC	ACC	ACC	ATC	TGG	CAG
114.	FGS114	CGCGCGC	ACATG	TCT	GCC	ACC	ACC	ATC	TGG	CAG	CAG
115.	FGS115	CGCGCGC	ACATG	TCT	ACC	ACC	ATC	TGG	CAG	CAG	ATG
116.	FGS116	CGCGCGC	ACATG	TCT	ACC	ATC	TGG	CAG	CAG	ATG	GAA
117.	FGS117	CGCGCGC	ACATG	TCT	ATC	TGG	CAG	CAG	ATG	GAA	GAA
118.	FGS118	CGCGCGC	ACATG	TCT	TGG	CAG	CAG	ATG	GAA	GAA	CTG
119.	FGS119	CGCGCGC	ACATG	TCT	CAG	CAG	ATG	GAA	GAA	CTG	GGA
120.	FGS120	CGCGCGC	ACATG	TCT	CAG	ATG	GAA	GAA	CTG	GGA	ATG
121.	FGS121	CGCGCGC	ACATG	TCT	ATG	GAA	GAA	CTG	GGA	ATG	GCC
122.	FGS122	CGCGCGC	ACATG	TCT	GAA	GAA	CTG	GGA	ATG	GCC	CCT
123.	FGS123	CGCGCGC	ACATG	TCT	GAA	CTG	GGA	ATG	GCC	CCT	GCC
124.	FGS124	CGCGCGC	ACATG	TCT	CTG	GGA	ATG	GCC	CCT	GCC	CTG
125.	FGS125	CGCGCGC	ACATG	TCT	GGA	ATG	GCC	CCT	GCC	CTG	CAG
126.	FGS126	CGCGCGC	ACATG	TCT	ATG	GCC	CCT	GCC	CTG	CAG	CCC
127.	FGS127	CGCGCGC	ACATG	TCT	GCC	CCT	GCC	CTG	CAG	CCC	ACC
128.	FGS128	CGCGCGC	ACATG	TCT	CCT	GCC	CTG	CAG	CCC	ACC	CAG
129.	FGS129	CGCGCGC	ACATG	TCT	GCC	CTG	CAG	CCC	ACC	CAG	GGT
130.	FGS130	CGCGCGC	ACATG	TCT	CTG	CAG	CCC	ACC	CAG	GGT	GCC
131.	FGS131	CGCGCGC	ACATG	TCT	CAG	CCC	ACC	CAG	GGT	GCC	ATG
132.	FGS132	CGCGCGC	ACATG	TCT	CCC	ACC	CAG	GGT	GCC	ATG	CCG
133.	FGS133	CGCGCGC	ACATG	TCT	ACC	CAG	GGT	GCC	ATG	CCG	GCC
134.	FGS134	CGCGCGC	ACATG	TCT	CAG	GGT	GCC	ATG	CCG	GCC	TTC
135.	FGS135	CGCGCGC	ACATG	TCT	GGT	GCC	ATG	CCG	GCC	TTC	GCC
136.	FGS136	CGCGCGC	ACATG	TCT	GCC	ATG	CCG	GCC	TTC	GCC	TCT
137.	FGS137	CGCGCGC	ACATG	TCT	ATG	CCG	GCC	TTC	GCC	TCT	GCT
138.	FGS138	CGCGCGC	ACATG	TCT	CCG	GCC	TTC	GCC	TCT	GCT	TTC
139.	FGS139	CGCGCGC	ACATG	TCT	GCC	TTC	GCC	TCT	GCT	TTC	CAG
140.	FGS140	CGCGCGC	ACATG	TCT	TTC	GCC	TCT	GCT	TTC	CAG	CGC
141.	FGS141	CGCGCGC	ACATG	TCT	GCC	TCT	GCT	TTC	CAG	CGC	CGG
142.	FGS142	CGCGCGC	ACATG	TCT	TCT	GCT	TTC	CAG	CGC	CGG	GCA
143.	FGS143	CGCGCGC	ACATG	TCT	GCT	TTC	CAG	CGC	CGG	GCA	GGA
144.	FGS144	CGCGCGC	ACATG	TCT	TTC	CAG	CGC	CGG	GCA	GGA	GGG
145.	FGS145	CGCGCGC	ACATG	TCT	CAG	CGC	CGG	GCA	GGA	GGG	GTC
146.	FGS146	CGCGCGC	ACATG	TCT	CGC	CGG	GCA	GGA	GGG	GTC	CTG
147.	FGS147	CGCGCGC	ACATG	TCT	CGG	GCA	GGA	GGG	GTC	CTG	GTT
148.	FGS148	CGCGCGC	ACATG	TCT	GCA	GGA	GGG	GTC	CTG	GTT	GCT
149.	FGS149	CGCGCGC	ACATG	TCT	GGA	GGG	GTC	CTG	GTT	GCT	AGC
150.	FGS150	CGCGCGC	ACATG	TCT	GGG	GTC	CTG	GTT	GCT	AGC	CAT
151.	FGS151	CGCGCGC	ACATG	TCT	GTC	CTG	GTT	GCT	AGC	CAT	CTG
152.	FGS152	CGCGCGC	ACATG	TCT	CTG	GTT	GCT	AGC	CAT	CTG	CAG
153.	FGS153	CGCGCGC	ACATG	TCT	GTT	GCT	AGC	CAT	CTG	CAG	AGC
154.	FGS154	CGCGCGC	ACATG	TCT	GCT	AGC	CAT	CTG	CAG	AGC	TTC
155.	FGS155	CGCGCGC	ACATG	TCT	AGC	CAT	CTG	CAG	AGC	TTC	CTG
156.	FGS156	CGCGCGC	ACATG	TCT	CAT	CTG	CAG	AGC	TTC	CTG	GAG
157.	FGS157	CGCGCGC	ACATG	TCT	CTG	CAG	AGC	TTC	CTG	GAG	GTG
158.	FGS158	CGCGCGC	ACATG	TCT	CAG	AGC	TTC	CTG	GAG	GTG	TCG
159.	FGS159	CGCGCGC	ACATG	TCT	AGC	TTC	CTG	GAG	GTG	TCG	TAC
160.	FGS160	CGCGCGC	ACATG	TCT	TTC	CTG	GAG	GTG	TCG	TAC	CGC
161.	FGS161	CGCGCGC	ACATG	TCT	CTG	GAG	GTG	TCG	TAC	CGC	GTT
162.	FGS162	CGCGCGC	ACATG	TCT	GAG	GTG	TCG	TAC	CGC	GTT	CTA
163.	FGS163	CGCGCGC	ACATG	TCT	GTG	TCG	TAC	CGC	GTT	CTA	CGC
164.	FGS164	CGCGCGC	ACATG	TCT	TCG	TAC	CGC	GTT	CTA	CGC	CAC
165.	FGS165	CGCGCGC	ACATG	TCT	TAC	CGC	GTT	CTA	CGC	CAC	GTT

166.	FGS166	CGCGCGC	ACATG	TCT	CGC	GTT	CTA	CGC	CAC	CTT	GCG
167.	FGS167	CGCGCGC	ACATG	TCT	GTT	CTA	CGC	CAC	CTT	GCG	CAG
168.	FGS168	CGCGCGC	ACATG	TCT	CTA	CGC	CAC	CTT	GCG	CAG	CCC
169.	FGS169A	CGCGCGC	ACATG	TCT	CGC	CAC	CTT	GCG	CAG	CCC	GA'C
170.	FGS170A	CGCGCGC	ACATG	TCT	CAC	CTT	GCG	CAG	CCC	GA'C	ATG
171.	FGS171A	CGCGCGC	ACATG	TCT	CTT	GCG	CAG	CCC	GA'C	ATG	GCT
172.	FGS172A	CGCGCGC	ACATG	TCT	GCG	CAG	CCC	GA'C	ATG	GCT	ACA
173.	FGS173A	CGCGCGC	ACATG	TCT	CAG	CCC	GA'C	ATG	GCT	ACA	CCA
174.	FGS174A	CGCGCGC	ACATG	TCT	CCC	GA'C	ATG	GCT	ACA	CCA	TTG
175.	FGS169B	CGCGCGC	ACATG	TCT	CGC	CAC	CTT	GCG	CAG	CCC	A'CT
176.	FGS170B	CGCGCGC	ACATG	TCT	CAC	CTT	GCG	CAG	CCC	A'CT	AGT
177.	FGS171B	CGCGCGC	ACATG	TCT	CTT	GCG	CAG	CCC	A'CT	AGT	CAT
178.	FGS172B	CGCGCGC	ACATG	TCT	GCG	CAG	CCC	A'CT	AGT	CAT	CCA
179.	FGS173B	CGCGCGC	ACATG	TCT	CAG	CCC	A'CT	AGT	CAT	CCA	CCT
180.	FGS174B	CGCGCGC	ACATG	TCT	CCC	A'CT	AGT	CAT	CCA	CCT	ATG
181.	FGS169C	CGCGCGC	ACATG	TCT	CGC	CAC	CTT	GCG	CAG	CCC	GGC
182.	FGS170C	CGCGCGC	ACATG	TCT	CAC	CTT	GCG	CAG	CCC	GGC	GGC
183.	FGS171C	CGCGCGC	ACATG	TCT	CTT	GCG	CAG	CCC	GGC	GGC	GGC
184.	FGS172C	CGCGCGC	ACATG	TCT	GCG	CAG	CCC	GGC	GGC	GGC	TCT
185.	FGS173C	CGCGCGC	ACATG	TCT	CAG	CCC	GGC	GGC	GGC	TCT	GA'C
186.	FGS174C	CGCGCGC	ACATG	TCT	CCC	GGC	GGC	GGC	TCT	GA'C	ATG
187.	RGS0A	TATATAT	GCGGCCGC	AGC	CAT	GTC	GGG	CTG	CGC	AAG	
188.	RGS0B	TATATAT	GCGGCCGC	AGC	CAT	GTC	ACG	CGT	ACG	ATT	
189.	RGS0C	TATATAT	GCGGCCGC	AGC	CAT	GTC	AGA	GCC	GCC	GCC	
190.	RGS1A	TATATAT	GCGGCCGC	TGT	AGC	CAT	GTC	GGG	CTG	CGC	
191.	RGS1B	TATATAT	GCGGCCGC	TGT	AGC	CAT	GTC	ACG	CGT	ACG	
192.	RGS1C	TATATAT	GCGGCCGC	TGT	AGC	CAT	GTC	AGA	GCC	GCC	
193.	RGS2A	TATATAT	GCGGCCGC	TGG	TGT	AGC	CAT	GTC	GGG	CTG	
194.	RGS2B	TATATAT	GCGGCCGC	TGG	TGT	AGC	CAT	GTC	ACG	CGT	
195.	RGS2C	TATATAT	GCGGCCGC	TGG	TGT	AGC	CAT	GTC	AGA	GCC	
196.	RGS3A	TATATAT	GCGGCCGC	CAA	TGG	TGT	AGC	CAT	GTC	GGG	
197.	RGS3B	TATATAT	GCGGCCGC	CAA	TGG	TGT	AGC	CAT	GTC	ACG	
198.	RGS3C	TATATAT	GCGGCCGC	CAA	TGG	TGT	AGC	CAT	GTC	AGA	
199.	RGS4	TATATAT	GCGGCCGC	GCC	CAA	TGG	TGT	AGC	CAT	GTC	
200.	RGS5	TATATAT	GCGGCCGC	AGG	GCC	CAA	TGG	TGT	AGC	CAT	
201.	RGS6	TATATAT	GCGGCCGC	GGC	AGG	GCC	CAA	TGG	TGT	AGC	
202.	RGS7	TATATAT	GCGGCCGC	GCT	GGC	AGG	GCC	CAA	TGG	TGT	
203.	RGS8	TATATAT	GCGGCCGC	GGA	GCT	GGC	AGG	GCC	CAA	TGG	
204.	RGS9	TATATAT	GCGGCCGC	CAG	GGA	GCT	GGC	AGG	GCC	CAA	
205.	RGS10	TATATAT	GCGGCCGC	GGG	CAG	GGA	GCT	GGC	AGG	GCC	
206.	RGS11	TATATAT	GCGGCCGC	CTG	GGG	CAG	GGA	GCT	GGC	AGG	
207.	RGS12	TATATAT	GCGGCCGC	GCT	CTG	GGG	CAG	GGA	GCT	GGC	
208.	RGS13	TATATAT	GCGGCCGC	GAA	GCT	CTG	GGG	CAG	GGA	GCT	
209.	RGS14	TATATAT	GCGGCCGC	CAG	GAA	GCT	CTG	GGG	CAG	GGA	
210.	RGS15	TATATAT	GCGGCCGC	GAG	CAG	GAA	GCT	CTG	GGG	CAG	
211.	RGS16	TATATAT	GCGGCCGC	CTT	GAG	CAG	GAA	GCT	CTG	GGG	
212.	RGS17	TATATAT	GCGGCCGC	AGA	CTT	GAG	CAG	GAA	GCT	CTG	
213.	RGS18	TATATAT	GCGGCCGC	TAA	AGA	CTT	GAG	CAG	GAA	GCT	
214.	RGS19	TATATAT	GCGGCCGC	CTC	TAA	AGA	CTT	GAG	CAG	GAA	
215.	RGS20	TATATAT	GCGGCCGC	TTG	CTC	TAA	AGA	CTT	GAG	CAG	
216.	RGS21	TATATAT	GCGGCCGC	CAC	TTG	CTC	TAA	AGA	CTT	GAG	
217.	RGS22	TATATAT	GCGGCCGC	CCT	CAC	TTG	CTC	TAA	AGA	CTT	
218.	RGS23	TATATAT	GCGGCCGC	CTT	CCT	CAC	TTG	CTC	TAA	AGA	
219.	RGS24	TATATAT	GCGGCCGC	GAT	CTT	CCT	CAC	TTG	CTC	TAA	
220.	RGS25	TATATAT	GCGGCCGC	CTG	GAT	CTT	CCT	CAC	TTG	CTC	
221.	RGS26	TATATAT	GCGGCCGC	GCC	CTG	GAT	CTT	CCT	CAC	TTG	
222.	RGS27	TATATAT	GCGGCCGC	ATC	GCC	CTG	GAT	CTT	CCT	CAC	
223.	RGS28	TATATAT	GCGGCCGC	GCC	ATC	GCC	CTG	GAT	CTT	CCT	
224.	RGS29	TATATAT	GCGGCCGC	TGC	GCC	ATC	GCC	CTG	GAT	CTT	
225.	RGS30	TATATAT	GCGGCCGC	CGC	TGC	GCC	ATC	GCC	CTG	GAT	
226.	RGS31	TATATAT	GCGGCCGC	GAG	CGC	TGC	GCC	ATC	GCC	CTG	
227.	RGS32	TATATAT	GCGGCCGC	CTG	GAG	CGC	TGC	GCC	ATC	GCC	
228.	RGS33	TATATAT	GCGGCCGC	CTC	CTG	GAG	CGC	TGC	GCC	ATC	
229.	RGS34	TATATAT	GCGGCCGC	CTT	CTC	CTG	GAG	CGC	TGC	GCC	
230.	RGS35	TATATAT	GCGGCCGC	CAG	CTT	CTC	CTG	GAG	CGC	TGC	
231.	RGS36	TATATAT	GCGGCCGC	ACA	CAG	CTT	CTC	CTG	GAG	CGC	
232.	RGS37	TATATAT	GCGGCCGC	GGC	ACA	CAG	CTT	CTC	CTG	GAG	
233.	RGS38	TATATAT	GCGGCCGC	GGT	GGC	ACA	CAG	CTT	CTC	CTG	
234.	RGS39	TATATAT	GCGGCCGC	GTA	GGT	GGC	ACA	CAG	CTT	CTC	
235.	RGS40	TATATAT	GCGGCCGC	CTT	GTA	GGT	GGC	ACA	CAG	CTT	
236.	RGS41	TATATAT	GCGGCCGC	CAG	CTT	GTA	GGT	GGC	ACA	CAG	
237.	RGS42	TATATAT	GCGGCCGC	GCA	CAG	CTT	GTA	GGT	GGC	ACA	
238.	RGS43	TATATAT	GCGGCCGC	GTG	GCA	CAG	CTT	GTA	GGT	GGC	
239.	RGS44	TATATAT	GCGGCCGC	GGG	GTG	GCA	CAG	CTT	GTA	GGT	
240.	RGS45	TATATAT	GCGGCCGC	CTC	GGG	GTG	GCA	CAG	CTT	GTA	
241.	RGS46	TATATAT	GCGGCCGC	CTC	CTC	GGG	GTG	GCA	CAG	CTT	
242.	RGS47	TATATAT	GCGGCCGC	CAG	CTC	CTC	GGG	GTG	GCA	CAG	
243.	RGS48	TATATAT	GCGGCCGC	CAC	CAG	CTC	CTC	GGG	GTG	GCA	
244.	RGS49	TATATAT	GCGGCCGC	CAG	CAC	CAG	CTC	CTC	GGG	GTG	
245.	RGS50	TATATAT	GCGGCCGC	GAG	CAG	CAC	CAG	CTC	CTC	GGG	
246.	RGS51	TATATAT	GCGGCCGC	TCC	GAG	CAG	CAC	CAG	CTC	CTC	
247.	RGS52	TATATAT	GCGGCCGC	GTG	TCC	GAG	CAG	CAC	CAG	CTC	
248.	RGS53	TATATAT	GCGGCCGC	AGA	GTG	TCC	GAG	CAG	CAC	CAG	
249.	RGS54	TATATAT	GCGGCCGC	CAG	AGA	GTG	TCC	GAG	CAG	CAC	
250.	RGS55	TATATAT	GCGGCCGC	GCC	CAG	AGA	GTG	TCC	GAG	CAG	
251.	RGS56	TATATAT	GCGGCCGC	GAT	GCC	CAG	AGA	GTG	TCC	GAG	

252.	RGS57	TATATAT	GCGGCCGC	GGG	GAT	GCC	CAG	AGA	GTG	TCC
253.	RGS58	TATATAT	GCGGCCGC	CCA	GGG	GAT	GCC	CAG	AGA	GTG
254.	RGS59	TATATAT	GCGGCCGC	AGC	CCA	GGG	GAT	GCC	CAG	AGA
255.	RGS60	TATATAT	GCGGCCGC	GGG	AGC	CCA	GGG	GAT	GCC	CAG
256.	RGS61	TATATAT	GCGGCCGC	CAG	GGG	AGC	CCA	GGG	GAT	GCC
257.	RGS62	TATATAT	GCGGCCGC	GCT	CAG	GGG	AGC	CCA	GGG	GAT
258.	RGS63	TATATAT	GCGGCCGC	GGA	GCT	CAG	GGG	AGC	CCA	GGG
259.	RGS64	TATATAT	GCGGCCGC	GCA	GGA	GCT	CAG	GGG	AGC	CCA
260.	RGS65	TATATAT	GCGGCCGC	GGG	GCA	GGA	GCT	CAG	GGG	AGC
261.	RGS66	TATATAT	GCGGCCGC	GCT	GGG	GCA	GGA	GCT	CAG	GGG
262.	RGS67	TATATAT	GCGGCCGC	CTG	GCT	GGG	GCA	GGA	GCT	CAG
263.	RGS68	TATATAT	GCGGCCGC	GGC	CTG	GCT	GGG	GCA	GGA	GCT
264.	RGS69	TATATAT	GCGGCCGC	CAG	GGC	CTG	GCT	GGG	GCA	GGA
265.	RGS70	TATATAT	GCGGCCGC	CTG	CAG	GGC	CTG	GCT	GGG	GCA
266.	RGS71	TATATAT	GCGGCCGC	CAG	CTG	CAG	GGC	CTG	GCT	GGG
267.	RGS72	TATATAT	GCGGCCGC	TGC	CAG	CTG	CAG	GGC	CTG	GCT
268.	RGS73	TATATAT	GCGGCCGC	GCC	TGC	CAG	CTG	CAG	GGC	CTG
269.	RGS74	TATATAT	GCGGCCGC	GCA	GCC	TGC	CAG	CTG	CAG	GGC
270.	RGS75	TATATAT	GCGGCCGC	CAA	GCA	GCC	TGC	CAG	CTG	CAG
271.	RGS76	TATATAT	GCGGCCGC	GCT	CAA	GCA	GCC	TGC	CAG	CTG
272.	RGS77	TATATAT	GCGGCCGC	GTT	GCT	CAA	GCA	GCC	TGC	CAG
273.	RGS78	TATATAT	GCGGCCGC	GAG	GTT	GCT	CAA	GCA	GCC	TGC
274.	RGS79	TATATAT	GCGGCCGC	ATG	GAG	GTT	GCT	CAA	GCA	GCC
275.	RGS80	TATATAT	GCGGCCGC	GCT	ATG	GAG	GTT	GCT	CAA	GCA
276.	RGS81	TATATAT	GCGGCCGC	GCC	GCT	ATG	GAG	GTT	GCT	CAA
277.	RGS82	TATATAT	GCGGCCGC	AAG	GCC	GCT	ATG	GAG	GTT	GCT
278.	RGS83	TATATAT	GCGGCCGC	GAA	AAG	GCC	GCT	ATG	GAG	GTT
279.	RGS84	TATATAT	GCGGCCGC	GAG	GAA	AAG	GCC	GCT	ATG	GAG
280.	RGS85	TATATAT	GCGGCCGC	GTA	GAG	GAA	AAG	GCC	GCT	ATG
281.	RGS86	TATATAT	GCGGCCGC	CTG	GTA	GAG	GAA	AAG	GCC	GCT
282.	RGS87	TATATAT	GCGGCCGC	CCC	CTG	GTA	GAG	GAA	AAG	GCC
283.	RGS88	TATATAT	GCGGCCGC	GAG	CCC	CTG	GTA	GAG	GAA	AAG
284.	RGS89	TATATAT	GCGGCCGC	CAG	GAG	CCC	CTG	GTA	GAG	GAA
285.	RGS90	TATATAT	GCGGCCGC	CTG	CAG	GAG	CCC	CTG	GTA	GAG
286.	RGS91	TATATAT	GCGGCCGC	GGC	CTG	CAG	GAG	CCC	CTG	GTA
287.	RGS92	TATATAT	GCGGCCGC	CAG	GGC	CTG	CAG	GAG	CCC	CTG
288.	RGS93	TATATAT	GCGGCCGC	TTC	CAG	GGC	CTG	CAG	GAG	CCC
289.	RGS94	TATATAT	GCGGCCGC	CCC	TTC	CAG	GGC	CTG	CAG	GAG
290.	RGS95	TATATAT	GCGGCCGC	TAT	CCC	TTC	CAG	GGC	CTG	CAG
291.	RGS96	TATATAT	GCGGCCGC	GGA	TAT	CCC	TTC	CAG	GGC	CTG
292.	RGS97	TATATAT	GCGGCCGC	GGG	GGA	TAT	CCC	TTC	CAG	GGC
293.	RGS98	TATATAT	GCGGCCGC	CTC	GGG	GGA	TAT	CCC	TTC	CAG
294.	RGS99	TATATAT	GCGGCCGC	CAA	CTC	GGG	GGA	TAT	CCC	TTC
295.	RGS100	TATATAT	GCGGCCGC	ACC	CAA	CTC	GGG	GGA	TAT	CCC
296.	RGS101	TATATAT	GCGGCCGC	GGG	ACC	CAA	CTC	GGG	GGA	TAT
297.	RGS102	TATATAT	GCGGCCGC	GGT	GGG	ACC	CAA	CTC	GGG	GGA
298.	RGS103	TATATAT	GCGGCCGC	CAA	GGT	GGG	ACC	CAA	CTC	GGG
299.	RGS104	TATATAT	GCGGCCGC	GTC	CAA	GGT	GGG	ACC	CAA	CTC
300.	RGS105	TATATAT	GCGGCCGC	TGT	GTC	CAA	GGT	GGG	ACC	CAA
301.	RGS106	TATATAT	GCGGCCGC	CAG	TGT	GTC	CAA	GGT	GGG	ACC
302.	RGS107	TATATAT	GCGGCCGC	CTG	CAG	TGT	GTC	CAA	GGT	GGG
303.	RGS108	TATATAT	GCGGCCGC	CAG	CTG	CAG	TGT	GTC	CAA	GGT
304.	RGS109	TATATAT	GCGGCCGC	GTC	CAG	CTG	CAG	TGT	GTC	CAA
305.	RGS110	TATATAT	GCGGCCGC	GAC	GTC	CAG	CTG	CAG	TGT	GTC
306.	RGS111	TATATAT	GCGGCCGC	GGC	GAC	GTC	CAG	CTG	CAG	TGT
307.	RGS112	TATATAT	GCGGCCGC	GTC	GGC	GAC	GTC	CAG	CTG	CAG
308.	RGS113	TATATAT	GCGGCCGC	AAA	GTC	GGC	GAC	GTC	CAG	CTG
309.	RGS114	TATATAT	GCGGCCGC	GGC	AAA	GTC	GGC	GAC	GTC	CAG
310.	RGS115	TATATAT	GCGGCCGC	GGT	GGC	AAA	GTC	GGC	GAC	GTC
311.	RGS116	TATATAT	GCGGCCGC	GGT	GGT	GGC	AAA	GTC	GGC	GAC
312.	RGS117	TATATAT	GCGGCCGC	GAT	GGT	GGT	GGC	AAA	GTC	GGC
313.	RGS118	TATATAT	GCGGCCGC	CCA	GAT	GGT	GGT	GGC	AAA	GTC
314.	RGS119	TATATAT	GCGGCCGC	CTG	CCA	GAT	GGT	GGT	GGC	AAA
315.	RGS120	TATATAT	GCGGCCGC	CTG	CTG	CCA	GAT	GGT	GGT	GGC
316.	RGS121	TATATAT	GCGGCCGC	CAT	CTG	CTG	CCA	GAT	GGT	GGT
317.	RGS122	TATATAT	GCGGCCGC	TTC	CAT	CTG	CTG	CCA	GAT	GGT
318.	RGS123	TATATAT	GCGGCCGC	TTC	TTC	CAT	CTG	CTG	CCA	GAT
319.	RGS124	TATATAT	GCGGCCGC	CAG	TTC	TTC	CAT	CTG	CTG	CCA
320.	RGS125	TATATAT	GCGGCCGC	TCC	CAG	TTC	TTC	CAT	CTG	CTG
321.	RGS126	TATATAT	GCGGCCGC	CAT	TCC	CAG	TTC	TTC	CAT	CTG
322.	RGS127	TATATAT	GCGGCCGC	GGC	CAT	TCC	CAG	TTC	TTC	CAT
323.	RGS128	TATATAT	GCGGCCGC	AGG	GGC	CAT	TCC	CAG	TTC	TTC
324.	RGS129	TATATAT	GCGGCCGC	GGC	AGG	GGC	CAT	TCC	CAG	TTC
325.	RGS130	TATATAT	GCGGCCGC	CAG	GGC	AGG	GGC	CAT	TCC	CAG
326.	RGS131	TATATAT	GCGGCCGC	CTG	CAG	GGC	AGG	GGC	CAT	TCC
327.	RGS132	TATATAT	GCGGCCGC	GGG	CTG	CAG	GGC	AGG	GGC	CAT
328.	RGS133	TATATAT	GCGGCCGC	GGT	GGG	CTG	CAG	GGC	AGG	GGC
329.	RGS134	TATATAT	GCGGCCGC	CTG	GGT	GGG	CTG	CAG	GGC	AGG
330.	RGS135	TATATAT	GCGGCCGC	ACC	CTG	GGT	GGG	CTG	CAG	GGC
331.	RGS136	TATATAT	GCGGCCGC	GGC	ACC	CTG	GGT	GGG	CTG	CAG
332.	RGS137	TATATAT	GCGGCCGC	CAT	GGC	ACC	CTG	GGT	GGG	CTG
333.	RGS138	TATATAT	GCGGCCGC	CGG	CAT	GGC	ACC	CTG	GGT	GGG
334.	RGS139	TATATAT	GCGGCCGC	GGC	CGG	CAT	GGC	ACC	CTG	GGT
335.	RGS140	TATATAT	GCGGCCGC	GAA	GGC	CGG	CAT	GGC	ACC	CTG
336.	RGS141	TATATAT	GCGGCCGC	GGC	GAA	GGC	CGG	CAT	GGC	ACC
337.	RGS142	TATATAT	GCGGCCGC	AGA	GGC	GAA	GGC	CGG	CAT	GGC

338.	RGS143	TATATAT	GCGGCCGC	AGC	AGA	GGC	GAA	GGC	CGG	CAT
339.	RGS144	TATATAT	GCGGCCGC	GAA	AGC	AGA	GGC	GAA	GGC	CGG
340.	RGS145	TATATAT	GCGGCCGC	CTG	GAA	AGC	AGA	GGC	GAA	GGC
341.	RGS146	TATATAT	GCGGCCGC	GCG	CTG	GAA	AGC	AGA	GGC	GAA
342.	RGS147	TATATAT	GCGGCCGC	CCG	GCG	CTG	GAA	AGC	AGA	GGC
343.	RGS148	TATATAT	GCGGCCGC	TGC	CCG	GCG	CTG	GAA	AGC	AGA
344.	RGS149	TATATAT	GCGGCCGC	TCC	TGC	CCG	GCG	CTG	GAA	AGC
345.	RGS150	TATATAT	GCGGCCGC	CCC	TCC	TGC	CCG	GCG	CTG	GAA
346.	RGS151	TATATAT	GCGGCCGC	GAC	CCC	TCC	TGC	CCG	GCG	CTG
347.	RGS152	TATATAT	GCGGCCGC	CAG	GAC	CCC	TCC	TGC	CCG	GCG
348.	RGS153	TATATAT	GCGGCCGC	AAC	CAG	GAC	CCC	TCC	TGC	CCG
349.	RGS154	TATATAT	GCGGCCGC	AGC	AAC	CAG	GAC	CCC	TCC	TGC
350.	RGS155	TATATAT	GCGGCCGC	GCT	AGC	AAC	CAG	GAC	CCC	TCC
351.	RGS156	TATATAT	GCGGCCGC	ATG	GCT	AGC	AAC	CAG	GAC	CCC
352.	RGS157	TATATAT	GCGGCCGC	CAG	ATG	GCT	AGC	AAC	CAG	GAC
353.	RGS158	TATATAT	GCGGCCGC	CTG	CAG	ATG	GCT	AGC	AAC	CAG
354.	RGS159	TATATAT	GCGGCCGC	GCT	CTG	CAG	ATG	GCT	AGC	AAC
355.	RGS160	TATATAT	GCGGCCGC	GAA	GCT	CTG	CAG	ATG	GCT	AGC
356.	RGS161	TATATAT	GCGGCCGC	CAG	GAA	GCT	CTG	CAG	ATG	GCT
357.	RGS162	TATATAT	GCGGCCGC	CTC	CAG	GAA	GCT	CTG	CAG	ATG
358.	RGS163	TATATAT	GCGGCCGC	CAC	CTC	CAG	GAA	GCT	CTG	CAG
359.	RGS164	TATATAT	GCGGCCGC	CGA	CAC	CTC	CAG	GAA	GCT	CTG
360.	RGS165	TATATAT	GCGGCCGC	GTA	CGA	CAC	CTC	CAG	GAA	GCT
361.	RGS166	TATATAT	GCGGCCGC	GCG	GTA	CGA	CAC	CTC	CAG	GAA
362.	RGS167	TATATAT	GCGGCCGC	AAC	GCG	GTA	CGA	CAC	CTC	CAG
363.	RGS168	TATATAT	GCGGCCGC	TAG	AAC	GCG	GTA	CGA	CAC	CTC
364.	RGS169	TATATAT	GCGGCCGC	GCG	TAG	AAC	GCG	GTA	CGA	CAC
365.	RGS170	TATATAT	GCGGCCGC	GTG	GCG	TAG	AAC	GCG	GTA	CGA
366.	RGS171	TATATAT	GCGGCCGC	AAG	GTG	GCG	TAG	AAC	GCG	GTA
367.	RGS172	TATATAT	GCGGCCGC	CGC	AAG	GTG	GCG	TAG	AAC	GCG
368.	RGS173	TATATAT	GCGGCCGC	CTG	CGC	AAG	GTG	GCG	TAG	AAC	GCG

CLAIMS

What is claimed is:

1. A method for making a biologically-active circularly-permuted protein of the formula C^1-L^1-N^1, derived from a parent protein of the formula N^1-C^1, wherein

C^1 is comprised of a segment derived from the carboxy portion of said parent protein;

N^1 is comprised of a segment derived from the amino terminal portion of said parent protein; and

L^1 is a chemical bond or a linker, linking C^1 to the amino terminus of L^1 and carboxy terminus of L^1 to the amino terminus of N^1;

comprising the steps of:

(a) making a series of circularly-permuted genes;

(b) inserting said circularly-permuted genes into a display vector;

(c) expressing said circularly-permuted genes such that the proteins encoded by said genes are presented on the surface of the display vector;

(d) generating a library of display vectors presenting the expressed circularly permuted protein;

(e) affinity-selecting the presenting display vectors with a target protein that can bind a biologically-active circularly-permuted protein;

(f) isolating and analyzing clones of selected display vectors to identify the presented circularly-permuted protein.

2. The method of claim 1 wherein the method of making a series of circularly-permuted genes is selected from the group consisting of making a tandemly-repeated intermediate, total synthesis of a synthetic gene, assembly of a gene from synthetic oligonucleotides, DNA amplification, and limited digestion of a circular

intermediate.

3. The method of claim 1 wherein said display vector is selected from the group consisting of bacteriophage display vectors, bacteria, and baculovirus vectors.

4. The method of claim 3, wherein said presentation vector is a bacteriophage.

5. The method of claim 4, wherein said presentation vector is bacteriophage M13.

6. The method of claim 5, wherein said presentation vector is a bacteriophage M13 gene III vector.

7. The method of claim 1 wherein said method of making a series of circularly permuted genes is a method of making a tandem repeat intermediate.

8 The method of claim 7 wherein said circularly-permuted genes are amplified from the repeat by gene amplification.

9. The method of claim 1 wherein said method of affinity selection comprises the steps consisting of:

(a) binding said presentation display vectors to a target protein;

(b) eluting said display vectors;

(c) amplifying said display vectors; and

(d) biopanning a pool of said amplified display vectors.

10. The method of claim 1 wherein L^1 is a linear peptide linker.

11. The method of claim 1 wherein said the DNA sequence encoding said linker L^1 is selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 368.

12. The method of claim 1 wherein the length of the C^1 in said permutein is shorter than the length of C^1 in said parent protein.

13. The method of claim 1 wherein the length of the N^1 in said permutein is shorter than the length of N^1 in said parent protein.

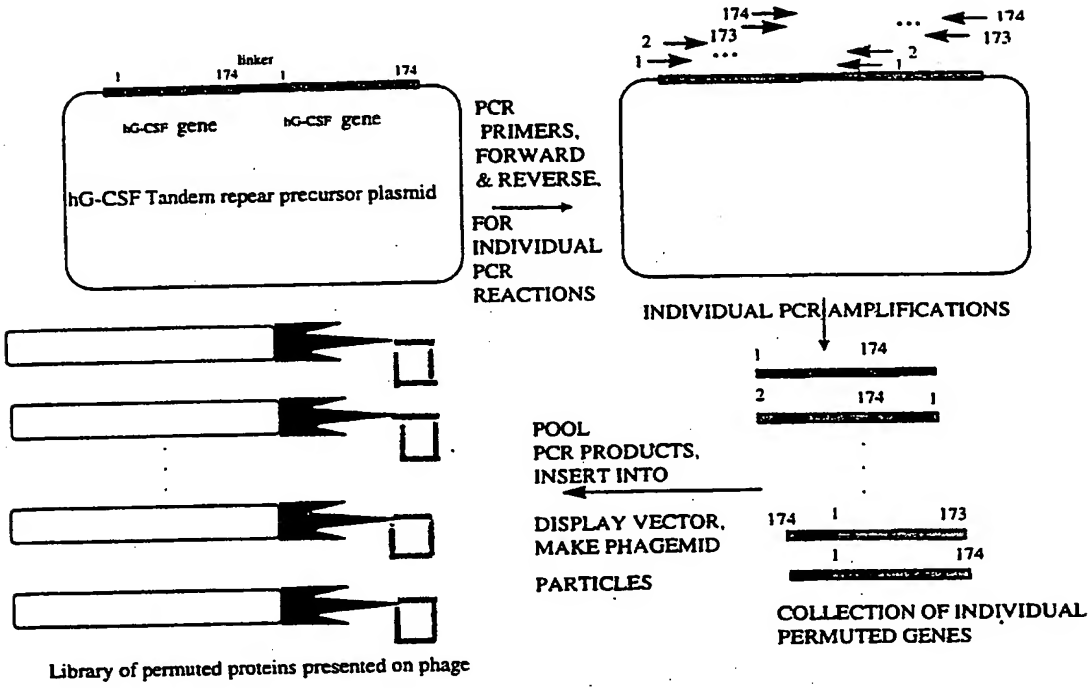
14. A circularly-permuted protein prepared by the method of claim 1.

15. A circularly-permuted protein of claim 14 comprising the G-CSF receptor agonist domain of a species of mylepoietin (MPO).

FIGURES

Figure 1A

1A. Constructing a scanning permutagenesis display library



5

Figure 1B

1B. Screening a display library

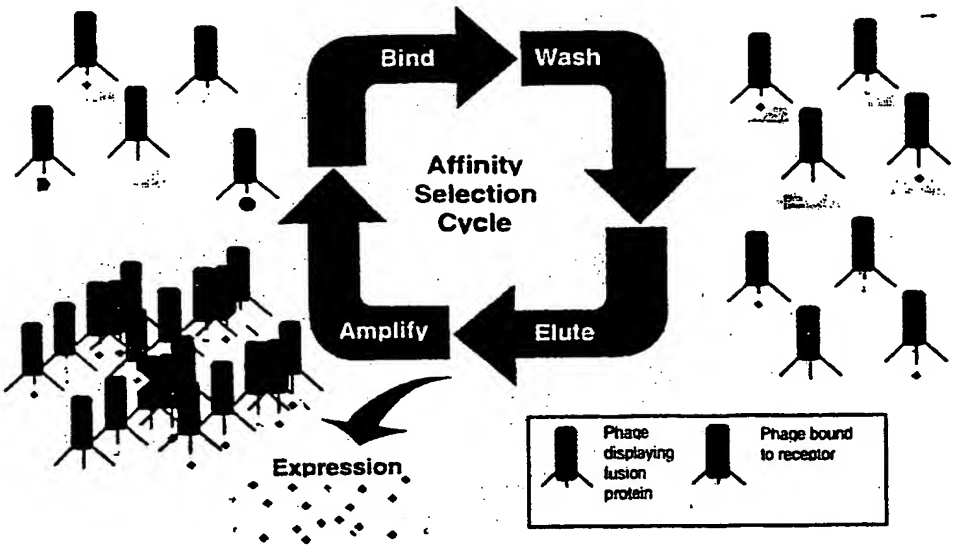


Figure 2

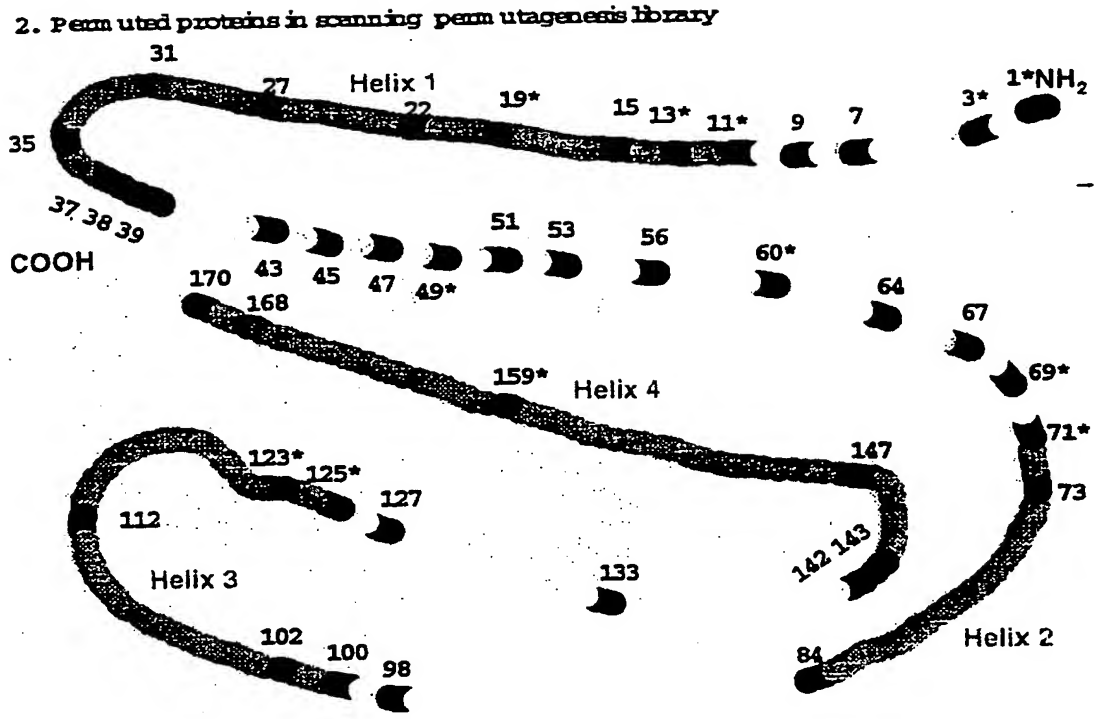
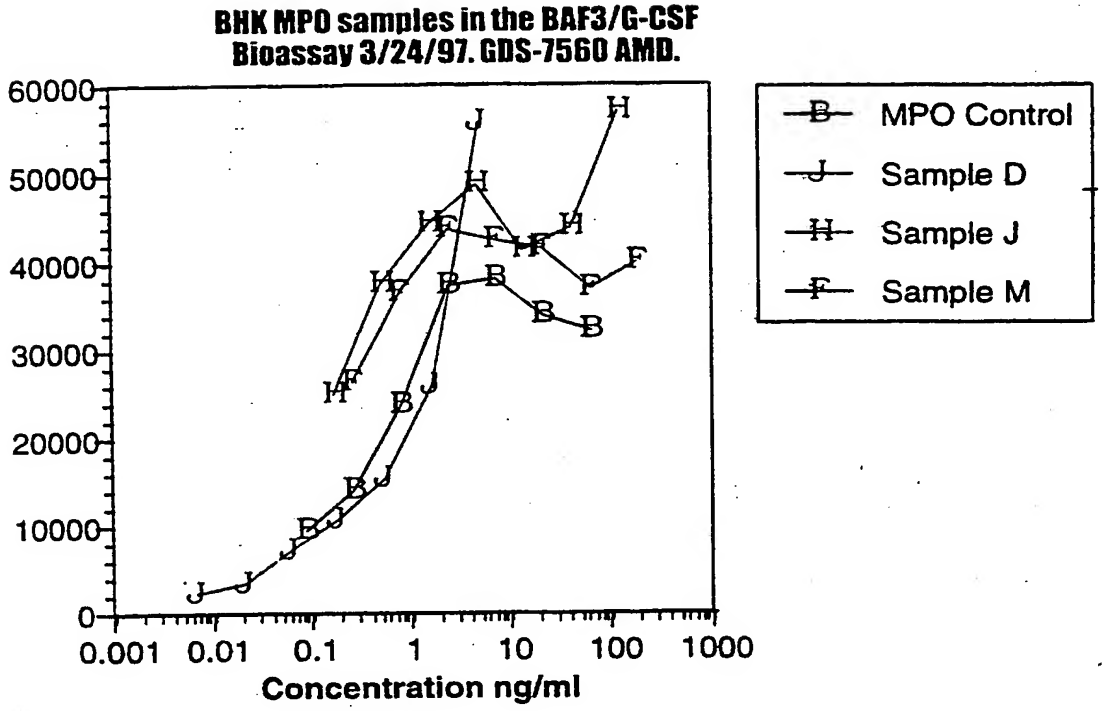

2. Permuted proteins in scanning permutagenesis library
31
27
Helix 1
22
19*
15
13*
11*
9
7
3*
1*NH2
35
37 38 39
COOH
51
53
56
60*
64
67
69*
71*
73
43
45
47
49*
170
168
159*
Helix 4
147
123*
125*
127
112
Helix 3
133
142 143
102
100
98
84
Helix 2

Figure 3


BHK MPO samples in the BAF3/G-CSF
Bioassay 3/24/97. GDS-7560 AMD.
60000
50000
40000
30000
20000
10000
0
0.001
0.01
0.1
1
10
100
1000
Concentration ng/ml
MPO Control
Sample D
Sample J
Sample M

SEQLIST.TXT

SEQUENCE LISTING

<110> Lee, Stephen C

<120> Method of Producing Permuteins by Scanning Permutagenesis

<130> C_3041

<140> 60/
<141> 1998-09-25

<160> 368

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 36
<212> DNA
<213> Synthetic

<400> 1
cgcgcgcaca tgtctacacc attgggccct gccagc 36

<210> 2
<211> 36
<212> DNA
<213> Synthetic

<400> 2
cgcgcgcaca tgtctccatt gggccctgcc agctcc 36

<210> 3
<211> 36
<212> DNA
<213> Synthetic

<400> 3
cgcgcgcaca tgtctttggg ccctgccagc tccctg 36

<210> 4
<211> 36
<212> DNA
<213> Synthetic

<400> 4
cgcgcgcaca tgtctggccc tgccagctcc ctgccc 36

<210> 5
<211> 36
<212> DNA
<213> Synthetic

<400> 5
cgcgcgcaca tgtctcctgc cagctccctg ccccag 36

<210> 6
<211> 36
<212> DNA
<213> Synthetic

<400> 6
cgcgcgcaca tgtctgccag ctcccctgccc cagagc 36

<210> 7
<211> 36
<212> DNA
<213> Synthetic

<400> 7
cgcgcgcaca tgtctagctc cctgccccag agcttc 36

<210> 8
<211> 36
<212> DNA
<213> Synthetic

<400> 8
cgcgcgcaca tgtcttccct gccccagagc ttcctg 36

<210> 9
<211> 36
<212> DNA
<213> Synthetic

<400> 9
cgcgcgcaca tgtctctgcc ccagagcttc ctgctc 36

<210> 10
<211> 36
<212> DNA
<213> Synthetic

<400> 10
cgcgcgcaca tgtctcccca gagcttcctg ctcaag 36

<210> 11
<211> 36
<212> DNA
<213> Synthetic

<400> 11
cgcgcgcaca tgtctcagag cttcctgctc aagtct 36

<210> 12
<211> 36
<212> DNA
<213> Synthetic

<400> 12
cgcgcgcaca tgtctagctt cctgctcaag tcttta 36

<210> 13
<211> 36
<212> DNA
<213> Synthetic

<400> 13
cgcgcgcaca tgtctttcct gctcaagtct ttagag 36

<210> 14
<211> 36
<212> DNA
<213> Synthetic

<400> 14
cgcgcgcaca tgtctctgct caagtcttta gagcaa 36

<210> 15
<211> 36
<212> DNA
<213> Synthetic

<400> 15
cgcgcgcaca tgtctctcaa gtctttagag caagtg 36

<210> 16
<211> 36
<212> DNA
<213> Synthetic

<400> 16
cgcgcgcaca tgtctaagtc tttagagcaa gtgagg 36

<210> 17
<211> 36
<212> DNA
<213> Synthetic

<400> 17
cgcgcgcaca tgtcttcttt agagcaagtg aggaag 36

<210> 18
<211> 36
<212> DNA
<213> Synthetic

<400> 18
cgcgcgcaca tgtctttaga gcaagtgagg aagatc 36

<210> 19
<211> 36
<212> DNA
<213> Synthetic

<400> 19
cgcgcgcaca tgtctgagca agtgaggaag atccag 36

<210> 20
<211> 36
<212> DNA
<213> Synthetic

<400> 20
cgcgcgcaca tgtctcaagt gaggaagatc cagggc 36

<210> 21
<211> 36
<212> DNA
<213> Synthetic

<400> 21
cgcgcgcaca tgtctgtgag gaagatccag ggcgat 36

<210> 22
<211> 36
<212> DNA
<213> Synthetic

<400> 22
cgcgcgcaca tgtctaggaa gatccagggc gatggc 36

<210> 23
<211> 36
<212> DNA
<213> Synthetic

<400> 23
cgcgcgcaca tgtctaagat ccagggcgat ggcgca 36

<210> 24

<211> 36
<212> DNA
<213> Synthetic

<400> 24
cgcgcgcaca tgtctatcca gggcgatggc gcagcg 36

<210> 25
<211> 36
<212> DNA
<213> Synthetic

<400> 25
cgcgcgcaca tgtctcaggg cgatggcgca gcgctc 36

<210> 26
<211> 36
<212> DNA
<213> Synthetic

<400> 26
cgcgcgcaca tgtctggcga tggcgcagcg ctccag 36

<210> 27
<211> 36
<212> DNA
<213> Synthetic

<400> 27
cgcgcgcaca tgtctgatgg cgcagcgctc caggag 36

<210> 28
<211> 36
<212> DNA
<213> Synthetic

<400> 28
cgcgcgcaca tgtctggcgc agcgctccag gagaag 36

<210> 29
<211> 36
<212> DNA
<213> Synthetic

<400> 29
cgcgcgcaca tgtctgcagc gctccaggag aagctg 36

<210> 30
<211> 36
<212> DNA
<213> Synthetic

<400> 30
cgcgcgcaca tgtctgcgct ccaggagaag ctgtgt 36

<210> 31
<211> 36
<212> DNA
<213> Synthetic

<400> 31
cgcgcgcaca tgtctctcca ggagaagctg tgtgcc 36

<210> 32
<211> 36
<212> DNA
<213> Synthetic

<400> 32

cgcgcgcaca tgtctcagga gaagctgtgt gccacc 36

<210> 33
<211> 36
<212> DNA
<213> Synthetic

<400> 33
cgcgcgcaca tgtctgagaa gctgtgtgcc acctac 36

<210> 34
<211> 36
<212> DNA
<213> Synthetic

<400> 34
cgcgcgcaca tgtctaagct gtgtgccacc tacaag 36

<210> 35
<211> 36
<212> DNA
<213> Synthetic

<400> 35
cgcgcgcaca tgtctctgtg tgccacctac aagctg 36

<210> 36
<211> 36
<212> DNA
<213> Synthetic

<400> 36
cgcgcgcaca tgtcttgtgc cacctacaag ctgtgc 36

<210> 37
<211> 36
<212> DNA
<213> Synthetic

<400> 37
cgcgcgcaca tgtctgccac ctacaagctg tgccac 36

<210> 38
<211> 36
<212> DNA
<213> Synthetic

<400> 38
cgcgcgcaca tgtctaccta caagctgtgc cacccc 36

<210> 39
<211> 36
<212> DNA
<213> Synthetic

<400> 39
cgcgcgcaca tgtcttacaa gctgtgccac cccgag 36

<210> 40
<211> 36
<212> DNA
<213> Synthetic

<400> 40
cgcgcgcaca tgtctaagct gtgccacccc gaggag 36

<210> 41
<211> 36
<212> DNA

<213> Synthetic

<400> 41
cgcgcgcaca tgtctctgtg ccacccccgag gagctg 36

<210> 42
<211> 36
<212> DNA
<213> Synthetic

<400> 42
cgcgcgcaca tgtcttgcca ccccgaggag ctggtg 36

<210> 43
<211> 36
<212> DNA
<213> Synthetic

<400> 43
cgcgcgcaca tgtctcaccc cgaggagctg gtgctg 36

<210> 44
<211> 36
<212> DNA
<213> Synthetic

<400> 44
cgcgcgcaca tgtctcccga ggagctggtg ctgctc 36

<210> 45
<211> 36
<212> DNA
<213> Synthetic

<400> 45
cgcgcgcaca tgtctgagga gctggtgctg ctcgga 36

<210> 46
<211> 36
<212> DNA
<213> Synthetic

<400> 46
cgcgcgcaca tgtctgagct ggtgctgctc ggacac 36

<210> 47
<211> 36
<212> DNA
<213> Synthetic

<400> 47
cgcgcgcaca tgtctctggt gctgctcgga cactct 36

<210> 48
<211> 36
<212> DNA
<213> Synthetic

<400> 48
cgcgcgcaca tgtctgtgct gctcggacac tctctg 36

<210> 49
<211> 36
<212> DNA
<213> Synthetic

<400> 49
cgcgcgcaca tgtctctgct cggacactct ctgggc 36

<210> 50
<211> 36
<212> DNA
<213> Synthetic

<400> 50
cgcgcgcaca tgtctctcgg acactctctg ggcatc 36

<210> 51
<211> 36
<212> DNA
<213> Synthetic

<400> 51
cgcgcgcaca tgtctggaca ctctctgggc atcccc 36

<210> 52
<211> 36
<212> DNA
<213> Synthetic

<400> 52
cgcgcgcaca tgtctcactc tctgggcatc ccctgg 36

<210> 53
<211> 36
<212> DNA
<213> Synthetic

<400> 53
cgcgcgcaca tgtcttctct gggcatcccc tgggct 36

<210> 54
<211> 36
<212> DNA
<213> Synthetic

<400> 54
cgcgcgcaca tgtctctggg catcccctgg gctccc 36

<210> 55
<211> 36
<212> DNA
<213> Synthetic

<400> 55
cgcgcgcaca tgtctggcat cccctgggct cccctg 36

<210> 56
<211> 36
<212> DNA
<213> Synthetic

<400> 56
cgcgcgcaca tgtctatccc ctgggctccc ctgagc 36

<210> 57
<211> 36
<212> DNA
<213> Synthetic

<400> 57
cgcgcgcaca tgtctccctg ggctcccctg agctcc 36

<210> 58
<211> 36
<212> DNA
<213> Synthetic

<400> 58
cgcgcgcaca tgtcttgggc tcccctgagc tcctgc 36

<210> 59
<211> 36
<212> DNA
<213> Synthetic

<400> 59
cgcgcgcaca tgtctgctcc cctgagctcc tgcccc 36

<210> 60
<211> 36
<212> DNA
<213> Synthetic

<400> 60
cgcgcgcaca tgtctccccc gagctcctgc cccagc 36

<210> 61
<211> 36
<212> DNA
<213> Synthetic

<400> 61
cgcgcgcaca tgtctctgag ctcctgcccc agccag 36

<210> 62
<211> 36
<212> DNA
<213> Synthetic

<400> 62
cgcgcgcaca tgtctagctc ctgccccagc caggcc 36

<210> 63
<211> 36
<212> DNA
<213> Synthetic

<400> 63
cgcgcgcaca tgtcttcctg ccccagccag gccctg 36

<210> 64
<211> 36
<212> DNA
<213> Synthetic

<400> 64
cgcgcgcaca tgtcttgccc cagccaggcc ctgcag 36

<210> 65
<211> 36
<212> DNA
<213> Synthetic

<400> 65
cgcgcgcaca tgtctcccag ccaggccctg cagctg 36

<210> 66
<211> 36
<212> DNA
<213> Synthetic

<400> 66
cgcgcgcaca tgtctagcca ggccctgcag ctggca 36

<210> 67
<211> 36

<212> DNA
<213> Synthetic

<400> 67
cgcgcgcaca tgtctcaggc cctgcagctg gcaggc 36

<210> 68
<211> 36
<212> DNA
<213> Synthetic

<400> 68
cgcgcgcaca tgtctgccct gcagctggca ggctgc 36

<210> 69
<211> 36
<212> DNA
<213> Synthetic

<400> 69
cgcgcgcaca tgtctctgca gctggcaggc tgcttg 36

<210> 70
<211> 36
<212> DNA
<213> Synthetic

<400> 70
cgcgcgcaca tgtctcagct ggcaggctgc ttgagc 36

<210> 71
<211> 36
<212> DNA
<213> Synthetic

<400> 71
cgcgcgcaca tgtctctggc aggctgcttg agccaa 36

<210> 72
<211> 36
<212> DNA
<213> Synthetic

<400> 72
cgcgcgcaca tgtctgcagg ctgcttgagc caactc 36

<210> 73
<211> 36
<212> DNA
<213> Synthetic

<400> 73
cgcgcgcaca tgtctggctg cttgagccaa ctccat 36

<210> 74
<211> 36
<212> DNA
<213> Synthetic

<400> 74
cgcgcgcaca tgtcttgctt gagccaactc catagc 36

<210> 75
<211> 36
<212> DNA
<213> Synthetic

<400> 75
cgcgcgcaca tgtctttgag ccaactccat agcggc 36

<210> 76
<211> 36
<212> DNA
<213> Synthetic

<400> 76
cgcgcgcaca tgtctagcca actccatagc ggcctt 36

<210> 77
<211> 36
<212> DNA
<213> Synthetic

<400> 77
cgcgcgcaca tgtctcaact ccatagcggc cttttc 36

<210> 78
<211> 36
<212> DNA
<213> Synthetic

<400> 78
cgcgcgcaca tgtctctcca tagcggcctt ttcctc 36

<210> 79
<211> 36
<212> DNA
<213> Synthetic

<400> 79
cgcgcgcaca tgtctcatag cggccttttc ctctac 36

<210> 80
<211> 36
<212> DNA
<213> Synthetic

<400> 80
cgcgcgcaca tgtctagcgg ccttttcctc taccag 36

<210> 81
<211> 36
<212> DNA
<213> Synthetic

<400> 81
cgcgcgcaca tgtctggcct ttcctctac cagggg 36

<210> 82
<211> 36
<212> DNA
<213> Synthetic

<400> 82
cgcgcgcaca tgtctctttt cctctaccag gggctc 36

<210> 83
<211> 36
<212> DNA
<213> Synthetic

<400> 83
cgcgcgcaca tgtctttcct ctaccagggg ctcctg 36

<210> 84
<211> 36
<212> DNA
<213> Synthetic

<400> 84
cgcgcgcaca tgtctctcta ccaggggctc ctgcag 36

<210> 85
<211> 36
<212> DNA
<213> Synthetic

<400> 85
cgcgcgcaca tgtcttacca ggggctcctg caggcc 36

<210> 86
<211> 36
<212> DNA
<213> Synthetic

<400> 86
cgcgcgcaca tgtctcaggg gctcctgcag gccctg 36

<210> 87
<211> 36
<212> DNA
<213> Synthetic

<400> 87
cgcgcgcaca tgtctgggct cctgcaggcc ctggaa 36

<210> 88
<211> 36
<212> DNA
<213> Synthetic

<400> 88
cgcgcgcaca tgtctctcct gcaggccctg gaaggg 36

<210> 89
<211> 36
<212> DNA
<213> Synthetic

<400> 89
cgcgcgcaca tgtctctgca ggccctggaa gggata 36

<210> 90
<211> 36
<212> DNA
<213> Synthetic

<400> 90
cgcgcgcaca tgtctcaggc cctggaaggg atatcc 36

<210> 91
<211> 36
<212> DNA
<213> Synthetic

<400> 91
cgcgcgcaca tgtctgccct ggaagggata tccccc 36

<210> 92
<211> 36
<212> DNA
<213> Synthetic

<400> 92
cgcgcgcaca tgtctctgga agggatatcc cccgag 36

<210> 93

<211> 36
<212> DNA
<213> Synthetic

<400> 93
cgcgcgcaca tgtctgaagg gatatccccc gagttg 36

<210> 94
<211> 36
<212> DNA
<213> Synthetic

<400> 94
cgcgcgcaca tgtctgggat atcccccgag ttgggt 36

<210> 95
<211> 36
<212> DNA
<213> Synthetic

<400> 95
cgcgcgcaca tgtctatatc ccccgagttg ggtccc 36

<210> 96
<211> 36
<212> DNA
<213> Synthetic

<400> 96
cgcgcgcaca tgtcttcccc cgagttgggt cccacc 36

<210> 97
<211> 36
<212> DNA
<213> Synthetic

<400> 97
cgcgcgcaca tgtctcccga gttgggtccc accttg 36

<210> 98
<211> 36
<212> DNA
<213> Synthetic

<400> 98
cgcgcgcaca tgtctgagtt gggtcccacc ttggac 36

<210> 99
<211> 36
<212> DNA
<213> Synthetic

<400> 99
cgcgcgcaca tgtctttggg tcccaccttg gacaca 36

<210> 100
<211> 36
<212> DNA
<213> Synthetic

<400> 100
cgcgcgcaca tgtctggtcc caccttggac acactg 36

<210> 101
<211> 36
<212> DNA
<213> Synthetic

<400> 101

cgcgcgcaca tgtctcccac cttggacaca ctgcag 36

<210> 102
<211> 36
<212> DNA
<213> Synthetic

<400> 102
cgcgcgcaca tgtctacctt ggacacactg cagctg 36

<210> 103
<211> 36
<212> DNA
<213> Synthetic

<400> 103
cgcgcgcaca tgtctttgga cacactgcag ctggac 36

<210> 104
<211> 36
<212> DNA
<213> Synthetic

<400> 104
cgcgcgcaca tgtctgacac actgcagctg gacgtc 36

<210> 105
<211> 36
<212> DNA
<213> Synthetic

<400> 105
cgcgcgcaca tgtctacact gcagctggac gtcgcc 36

<210> 106
<211> 36
<212> DNA
<213> Synthetic

<400> 106
cgcgcgcaca tgtctctgca gctggacgtc gccgac 36

<210> 107
<211> 36
<212> DNA
<213> Synthetic

<400> 107
cgcgcgcaca tgtctcagct ggacgtcgcc gacttt 36

<210> 108
<211> 36
<212> DNA
<213> Synthetic

<400> 108
cgcgcgcaca tgtctctgga cgtcgccgac tttgcc 36

<210> 109
<211> 36
<212> DNA
<213> Synthetic

<400> 109
cgcgcgcaca tgtctgacgt cgccgacttt gccacc 36

<210> 110
<211> 36
<212> DNA

<213> Synthetic

<400> 110
cgcgcgcaca tgtctgtcgc cgactttgcc accacc 36

<210> 111
<211> 36
<212> DNA
<213> Synthetic

<400> 111
cgcgcgcaca tgtctgccga ctttgccacc accatc 36

<210> 112
<211> 36
<212> DNA
<213> Synthetic

<400> 112
cgcgcgcaca tgtctgactt tgccaccacc atctgg 36

<210> 113
<211> 36
<212> DNA
<213> Synthetic

<400> 113
cgcgcgcaca tgtcttttgc caccaccatc tggcag 36

<210> 114
<211> 36
<212> DNA
<213> Synthetic

<400> 114
cgcgcgcaca tgtctgccac caccatctgg cagcag 36

<210> 115
<211> 36
<212> DNA
<213> Synthetic

<400> 115
cgcgcgcaca tgtctaccac catctggcag cagatg 36

<210> 116
<211> 36
<212> DNA
<213> Synthetic

<400> 116
cgcgcgcaca tgtctaccat ctggcagcag atggaa 36

<210> 117
<211> 36
<212> DNA
<213> Synthetic

<400> 117
cgcgcgcaca tgtctatctg gcagcagatg gaagaa 36

<210> 118
<211> 36
<212> DNA
<213> Synthetic

<400> 118
cgcgcgcaca tgtcttggca gcagatggaa gaactg 36

<210> 119
<211> 36
<212> DNA
<213> Synthetic

<400> 119
cgcgcgcaca tgtctcagca gatggaagaa ctggga 36

<210> 120
<211> 36
<212> DNA
<213> Synthetic

<400> 120
cgcgcgcaca tgtctcagat ggaagaactg ggaatg 36

<210> 121
<211> 36
<212> DNA
<213> Synthetic

<400> 121
cgcgcgcaca tgtctatgga agaactggga atggcc 36

<210> 122
<211> 36
<212> DNA
<213> Synthetic

<400> 122
cgcgcgcaca tgtctgaaga actgggaatg gcccct 36

<210> 123
<211> 36
<212> DNA
<213> Synthetic

<400> 123
cgcgcgcaca tgtctgaact gggaatggcc cctgcc 36

<210> 124
<211> 36
<212> DNA
<213> Synthetic

<400> 124
cgcgcgcaca tgtctctggg aatggcccct gccctg 36

<210> 125
<211> 36
<212> DNA
<213> Synthetic

<400> 125
cgcgcgcaca tgtctggaat ggcccctgcc ctgcag 36

<210> 126
<211> 36
<212> DNA
<213> Synthetic

<400> 126
cgcgcgcaca tgtctatggc ccctgccctg cagccc 36

<210> 127
<211> 36
<212> DNA
<213> Synthetic

<400> 127
cgcgcgcaca tgtctgcccc tgccctgcag cccacc 36

<210> 128
<211> 36
<212> DNA
<213> Synthetic

<400> 128
cgcgcgcaca tgtctcctgc cctgcagccc acccag 36

<210> 129
<211> 36
<212> DNA
<213> Synthetic

<400> 129
cgcgcgcaca tgtctgccct gcagcccacc cagggt 36

<210> 130
<211> 36
<212> DNA
<213> Synthetic

<400> 130
cgcgcgcaca tgtctctgca gcccacccag ggtgcc 36

<210> 131
<211> 36
<212> DNA
<213> Synthetic

<400> 131
cgcgcgcaca tgtctcagcc cacccagggt gccatg 36

<210> 132
<211> 36
<212> DNA
<213> Synthetic

<400> 132
cgcgcgcaca tgtctcccac ccagggtgcc atgccg 36

<210> 133
<211> 36
<212> DNA
<213> Synthetic

<400> 133
cgcgcgcaca tgtctaccca gggtgccatg ccggcc 36

<210> 134
<211> 36
<212> DNA
<213> Synthetic

<400> 134
cgcgcgcaca tgtctcaggg tgccatgccg gccttc 36

<210> 135
<211> 36
<212> DNA
<213> Synthetic

<400> 135
cgcgcgcaca tgtctggtgc catgccggcc ttcgcc 36

<210> 136
<211> 36

<212> DNA
<213> Synthetic

<400> 136
cgcgcgcaca tgtctgccat gccggccttc gcctct 36

<210> 137
<211> 36
<212> DNA
<213> Synthetic

<400> 137
cgcgcgcaca tgtctatgcc ggccttcgcc tctgct 36

<210> 138
<211> 36
<212> DNA
<213> Synthetic

<400> 138
cgcgcgcaca tgtctccggc cttcgcctct gctttc 36

<210> 139
<211> 36
<212> DNA
<213> Synthetic

<400> 139
cgcgcgcaca tgtctgcctt cgcctctgct ttccag 36

<210> 140
<211> 36
<212> DNA
<213> Synthetic

<400> 140
cgcgcgcaca tgtctttcgc ctctgctttc cagcgc 36

<210> 141
<211> 36
<212> DNA
<213> Synthetic

<400> 141
cgcgcgcaca tgtctgcctc tgctttccag cgccgg 36

<210> 142
<211> 36
<212> DNA
<213> Synthetic

<400> 142
cgcgcgcaca tgtcttctgc tttccagcgc cgggca 36

<210> 143
<211> 36
<212> DNA
<213> Synthetic

<400> 143
cgcgcgcaca tgtctgcttt ccagcgccgg gcagga 36

<210> 144
<211> 36
<212> DNA
<213> Synthetic

<400> 144
cgcgcgcaca tgtctttcca gcgccgggca ggaggg 36

<210> 145
<211> 36
<212> DNA
<213> Synthetic

<400> 145
cgcgcgcaca tgtctcagcg ccgggcagga ggggtc 36

<210> 146
<211> 36
<212> DNA
<213> Synthetic

<400> 146
cgcgcgcaca tgtctcgccg ggcaggaggg gtcctg 36

<210> 147
<211> 36
<212> DNA
<213> Synthetic

<400> 147
cgcgcgcaca tgtctcgggc aggaggggtc ctggtt 36

<210> 148
<211> 36
<212> DNA
<213> Synthetic

<400> 148
cgcgcgcaca tgtctgcagg aggggtcctg gttgct 36

<210> 149
<211> 36
<212> DNA
<213> Synthetic

<400> 149
cgcgcgcaca tgtctggagg ggtcctggtt gctagc 36

<210> 150
<211> 36
<212> DNA
<213> Synthetic

<400> 150
cgcgcgcaca tgtctggggt cctggttgct agccat 36

<210> 151
<211> 36
<212> DNA
<213> Synthetic

<400> 151
cgcgcgcaca tgtctgtcct ggttgctagc catctg 36

<210> 152
<211> 36
<212> DNA
<213> Synthetic

<400> 152
cgcgcgcaca tgtctctggt tgctagccat ctgcag 36

<210> 153
<211> 36
<212> DNA
<213> Synthetic

<400> 153
cgcgcgcaca tgtctgttgc tagccatctg cagagc 36

<210> 154
<211> 36
<212> DNA
<213> Synthetic

<400> 154
cgcgcgcaca tgtctgctag ccatctgcag agcttc 36

<210> 155
<211> 36
<212> DNA
<213> Synthetic

<400> 155
cgcgcgcaca tgtctagcca tctgcagagc ttcctg 36

<210> 156
<211> 36
<212> DNA
<213> Synthetic

<400> 156
cgcgcgcaca tgtctcatct gcagagcttc ctggag 36

<210> 157
<211> 36
<212> DNA
<213> Synthetic

<400> 157
cgcgcgcaca tgtctctgca gagcttcctg gaggtg 36

<210> 158
<211> 36
<212> DNA
<213> Synthetic

<400> 158
cgcgcgcaca tgtctcagag cttcctggag gtgtcg 36

<210> 159
<211> 36
<212> DNA
<213> Synthetic

<400> 159
cgcgcgcaca tgtctagctt cctggaggtg tcgtac 36

<210> 160
<211> 36
<212> DNA
<213> Synthetic

<400> 160
cgcgcgcaca tgtctttcct ggaggtgtcg taccgc 36

<210> 161
<211> 36
<212> DNA
<213> Synthetic

<400> 161
cgcgcgcaca tgtctctgga ggtgtcgtac cgcgtt 36

<210> 162

<211> 36
<212> DNA
<213> Synthetic

<400> 162
cgcgcgcaca tgtctgaggt gtcgtaccgc gttcta 36

<210> 163
<211> 36
<212> DNA
<213> Synthetic

<400> 163
cgcgcgcaca tgtctgtgtc gtaccgcgtt ctacgc 36

<210> 164
<211> 36
<212> DNA
<213> Synthetic

<400> 164
cgcgcgcaca tgtcttcgta ccgcgttcta cgccac 36

<210> 165
<211> 36
<212> DNA
<213> Synthetic

<400> 165
cgcgcgcaca tgtcttaccg cgttctacgc cacctt 36

<210> 166
<211> 36
<212> DNA
<213> Synthetic

<400> 166
cgcgcgcaca tgtctcgcgt tctacgccac cttgcg 36

<210> 167
<211> 36
<212> DNA
<213> Synthetic

<400> 167
cgcgcgcaca tgtctgttct acgccacctt gcgcag 36

<210> 168
<211> 36
<212> DNA
<213> Synthetic

<400> 168
cgcgcgcaca tgtctctacg ccaccttgcg cagccc 36

<210> 169
<211> 36
<212> DNA
<213> Synthetic

<400> 169
cgcgcgcaca tgtctcgcca ccttgcgcag cccgac 36

<210> 170
<211> 36
<212> DNA
<213> Synthetic

<400> 170

cgcgcgcaca tgtctcacct tgcgcagccc gacatg 36

<210> 171
<211> 36
<212> DNA
<213> Synthetic

<400> 171
cgcgcgcaca tgtctcttgc gcagcccgac atggct 36

<210> 172
<211> 36
<212> DNA
<213> Synthetic

<400> 172
cgcgcgcaca tgtctgcgca gcccgacatg gctaca 36

<210> 173
<211> 36
<212> DNA
<213> Synthetic

<400> 173
cgcgcgcaca tgtctcagcc cgacatggct acacca 36

<210> 174
<211> 36
<212> DNA
<213> Synthetic

<400> 174
cgcgcgcaca tgtctcccga catggctaca ccattg 36

<210> 175
<211> 36
<212> DNA
<213> Synthetic

<400> 175
cgcgcgcaca tgtctcgcca ccttgcgcag cccact 36

<210> 176
<211> 36
<212> DNA
<213> Synthetic

<400> 176
cgcgcgcaca tgtctcacct tgcgcagccc actagt 36

<210> 177
<211> 36
<212> DNA
<213> Synthetic

<400> 177
cgcgcgcaca tgtctcttgc gcagcccact agtcat 36

<210> 178
<211> 36
<212> DNA
<213> Synthetic

<400> 178
cgcgcgcaca tgtctgcgca gcccactagt catcca 36

<210> 179
<211> 36
<212> DNA

<213> Synthetic

<400> 179
cgcgcgcaca tgtctcagcc cactagtcat ccacct 36

<210> 180
<211> 36
<212> DNA
<213> Synthetic

<400> 180
cgcgcgcaca tgtctcccac tagtcatcca cctatg 36

<210> 181
<211> 36
<212> DNA
<213> Synthetic

<400> 181
cgcgcgcaca tgtctcgcca ccttgcgcag cccggc 36

<210> 182
<211> 36
<212> DNA
<213> Synthetic

<400> 182
cgcgcgcaca tgtctcacct tgcgcagccc ggcggc 36

<210> 183
<211> 36
<212> DNA
<213> Synthetic

<400> 183
cgcgcgcaca tgtctcttgc gcagcccggc ggcggc 36

<210> 184
<211> 36
<212> DNA
<213> Synthetic

<400> 184
cgcgcgcaca tgtctgcgca gcccggcggc ggctct 36

<210> 185
<211> 36
<212> DNA
<213> Synthetic

<400> 185
cgcgcgcaca tgtctcagcc cggcggcggc tctgac 36

<210> 186
<211> 36
<212> DNA
<213> Synthetic

<400> 186
cgcgcgcaca tgtctcccgg cggcggctct gacatg 36

<210> 187
<211> 36
<212> DNA
<213> Synthetic

<400> 187
tatatatgcg gccgcagcca tgtcgggctg cgcaag 36

<210> 188
<211> 36
<212> DNA
<213> Synthetic

<400> 188
tatatatgcg gccgcagcca tgtcacgcgt acgatt 36

<210> 189
<211> 36
<212> DNA
<213> Synthetic

<400> 189
tatatatgcg gccgcagcca tgtcagagcc gccgcc 36

<210> 190
<211> 36
<212> DNA
<213> Synthetic

<400> 190
tatatatgcg gccgctgtag ccatgtcggg ctgcgc 36

<210> 191
<211> 36
<212> DNA
<213> Synthetic

<400> 191
tatatatgcg gccgctgtag ccatgtcacg cgtacg 36

<210> 192
<211> 36
<212> DNA
<213> Synthetic

<400> 192
tatatatgcg gccgctgtag ccatgtcaga gccgcc 36

<210> 193
<211> 36
<212> DNA
<213> Synthetic

<400> 193
tatatatgcg gccgctggtg tagccatgtc gggctg 36

<210> 194
<211> 36
<212> DNA
<213> Synthetic

<400> 194
tatatatgcg gccgctggtg tagccatgtc acgcgt 36

<210> 195
<211> 36
<212> DNA
<213> Synthetic

<400> 195
tatatatgcg gccgctggtg tagccatgtc agagcc 36

<210> 196
<211> 36
<212> DNA
<213> Synthetic

<400> 196
tatatatgcg gccgccaatg gtgtagccat gtcggg 36

<210> 197
<211> 36
<212> DNA
<213> Synthetic

<400> 197
tatatatgcg gccgccaatg gtgtagccat gtcacg 36

<210> 198
<211> 36
<212> DNA
<213> Synthetic

<400> 198
tatatatgcg gccgccaatg gtgtagccat gtcaga 36

<210> 199
<211> 36
<212> DNA
<213> Synthetic

<400> 199
tatatatgcg gccgcgccca atggtgtagc catgtc 36

<210> 200
<211> 36
<212> DNA
<213> Synthetic

<400> 200
tatatatgcg gccgcagggc ccaatggtgt agccat 36

<210> 201
<211> 36
<212> DNA
<213> Synthetic

<400> 201
tatatatgcg gccgcggcag ggcccaatgg tgtagc 36

<210> 202
<211> 36
<212> DNA
<213> Synthetic

<400> 202
tatatatgcg gccgcgctgg cagggcccaa tggtgt 36

<210> 203
<211> 36
<212> DNA
<213> Synthetic

<400> 203
tatatatgcg gccgcggagc tggcagggcc caatgg 36

<210> 204
<211> 36
<212> DNA
<213> Synthetic

<400> 204
tatatatgcg gccgccaggg agctggcagg gcccaa 36

<210> 205
<211> 36

<212> DNA
<213> Synthetic

<400> 205
tatatatgcg gccgcgggca gggagctggc agggcc 36

<210> 206
<211> 36
<212> DNA
<213> Synthetic

<400> 206
tatatatgcg gccgcctggg gcagggagct ggcagg 36

<210> 207
<211> 36
<212> DNA
<213> Synthetic

<400> 207
tatatatgcg gccgcgctct ggggcaggga gctggc 36

<210> 208
<211> 36
<212> DNA
<213> Synthetic

<400> 208
tatatatgcg gccgcgaagc tctggggcag ggagct 36

<210> 209
<211> 36
<212> DNA
<213> Synthetic

<400> 209
tatatatgcg gccgccagga agctctgggg caggga 36

<210> 210
<211> 36
<212> DNA
<213> Synthetic

<400> 210
tatatatgcg gccgcgagca ggaagctctg gggcag 36

<210> 211
<211> 36
<212> DNA
<213> Synthetic

<400> 211
tatatatgcg gccgccttga gcaggaagct ctgggg 36

<210> 212
<211> 36
<212> DNA
<213> Synthetic

<400> 212
tatatatgcg gccgcagact tgagcaggaa gctctg 36

<210> 213
<211> 36
<212> DNA
<213> Synthetic

<400> 213
tatatatgcg gccgctaaag acttgagcag gaagct 36

<210> 214
<211> 36
<212> DNA
<213> Synthetic

<400> 214
tatatatgcg gccgcctcta aagacttgag caggaa 36

<210> 215
<211> 36
<212> DNA
<213> Synthetic

<400> 215
tatatatgcg gccgcttgct ctaaagactt gagcag 36

<210> 216
<211> 36
<212> DNA
<213> Synthetic

<400> 216
tatatatgcg gccgccactt gctctaaaga cttgag 36

<210> 217
<211> 36
<212> DNA
<213> Synthetic

<400> 217
tatatatgcg gccgccctca cttgctctaa agactt 36

<210> 218
<211> 36
<212> DNA
<213> Synthetic

<400> 218
tatatatgcg gccgccttcc tcacttgctc taaaga 36

<210> 219
<211> 36
<212> DNA
<213> Synthetic

<400> 219
tatatatgcg gccgcgatct tcctcacttg ctctaa 36

<210> 220
<211> 36
<212> DNA
<213> Synthetic

<400> 220
tatatatgcg gccgcctgga tcttcctcac ttgctc 36

<210> 221
<211> 36
<212> DNA
<213> Synthetic

<400> 221
tatatatgcg gccgcgccct ggatcttcct cacttg 36

<210> 222
<211> 36
<212> DNA
<213> Synthetic

<400> 222
tatatatgcg gccgcatcgc cctggatctt cctcac 36

<210> 223
<211> 36
<212> DNA
<213> Synthetic

<400> 223
tatatatgcg gccgcgccat cgccctggat cttcct 36

<210> 224
<211> 36
<212> DNA
<213> Synthetic

<400> 224
tatatatgcg gccgctgcgc catcgccctg gatctt 36

<210> 225
<211> 36
<212> DNA
<213> Synthetic

<400> 225
tatatatgcg gccgccgctg cgccatcgcc ctggat 36

<210> 226
<211> 36
<212> DNA
<213> Synthetic

<400> 226
tatatatgcg gccgcgagcg ctgcgccatc gccctg 36

<210> 227
<211> 36
<212> DNA
<213> Synthetic

<400> 227
tatatatgcg gccgcctgga gcgctgcgcc atcgcc 36

<210> 228
<211> 36
<212> DNA
<213> Synthetic

<400> 228
tatatatgcg gccgcctcct ggagcgctgc gccatc 36

<210> 229
<211> 36
<212> DNA
<213> Synthetic

<400> 229
tatatatgcg gccgccttct cctggagcgc tgcgcc 36

<210> 230
<211> 36
<212> DNA
<213> Synthetic

<400> 230
tatatatgcg gccgccagct tctcctggag cgctgc 36

<210> 231

<211> 36
<212> DNA
<213> Synthetic

<400> 231
tatatatgcg gccgcacaca gcttctcctg gagcgc 36

<210> 232
<211> 36
<212> DNA
<213> Synthetic

<400> 232
tatatatgcg gccgcggcac acagcttctc ctggag 36

<210> 233
<211> 36
<212> DNA
<213> Synthetic

<400> 233
tatatatgcg gccgcggtgg cacacagctt ctcctg 36

<210> 234
<211> 36
<212> DNA
<213> Synthetic

<400> 234
tatatatgcg gccgcgtagg tggcacacag cttctc 36

<210> 235
<211> 36
<212> DNA
<213> Synthetic

<400> 235
tatatatgcg gccgccttgt aggtggcaca cagctt 36

<210> 236
<211> 36
<212> DNA
<213> Synthetic

<400> 236
tatatatgcg gccgccagct tgtaggtggc acacag 36

<210> 237
<211> 36
<212> DNA
<213> Synthetic

<400> 237
tatatatgcg gccgcgcaca gcttgtaggt ggcaca 36

<210> 238
<211> 36
<212> DNA
<213> Synthetic

<400> 238
tatatatgcg gccgcgtggc acagcttgta ggtggc 36

<210> 239
<211> 36
<212> DNA
<213> Synthetic

<400> 239

tatatatgcg gccgcggggt ggcacagctt gtaggt 36

<210> 240
<211> 36
<212> DNA
<213> Synthetic

<400> 240
tatatatgcg gccgcctcgg ggtggcacag cttgta 36

<210> 241
<211> 36
<212> DNA
<213> Synthetic

<400> 241
tatatatgcg gccgcctcct cggggtggca cagctt 36

<210> 242
<211> 36
<212> DNA
<213> Synthetic

<400> 242
tatatatgcg gccgccagct cctcggggtg gcacag 36

<210> 243
<211> 36
<212> DNA
<213> Synthetic

<400> 243
tatatatgcg gccgccacca gctcctcggg gtggca 36

<210> 244
<211> 36
<212> DNA
<213> Synthetic

<400> 244
tatatatgcg gccgccagca ccagctcctc ggggtg 36

<210> 245
<211> 36
<212> DNA
<213> Synthetic

<400> 245
tatatatgcg gccgcgagca gcaccagctc ctcggg 36

<210> 246
<211> 36
<212> DNA
<213> Synthetic

<400> 246
tatatatgcg gccgctccga gcagcaccag ctcctc 36

<210> 247
<211> 36
<212> DNA
<213> Synthetic

<400> 247
tatatatgcg gccgcgtgtc cgagcagcac cagctc 36

<210> 248
<211> 36
<212> DNA

<213> Synthetic

<400> 248
tatatatgcg gccgcagagt gtccgagcag caccag 36

<210> 249
<211> 36
<212> DNA
<213> Synthetic

<400> 249
tatatatgcg gccgccagag agtgtccgag cagcac 36

<210> 250
<211> 36
<212> DNA
<213> Synthetic

<400> 250
tatatatgcg gccgcgccca gagagtgtcc gagcag 36

<210> 251
<211> 36
<212> DNA
<213> Synthetic

<400> 251
tatatatgcg gccgcgatgc ccagagagtg tccgag 36

<210> 252
<211> 36
<212> DNA
<213> Synthetic

<400> 252
tatatatgcg gccgcgggga tgcccagaga gtgtcc 36

<210> 253
<211> 36
<212> DNA
<213> Synthetic

<400> 253
tatatatgcg gccgcccagg ggatgcccag agagtg 36

<210> 254
<211> 36
<212> DNA
<213> Synthetic

<400> 254
tatatatgcg gccgcagccc aggggatgcc cagaga 36

<210> 255
<211> 36
<212> DNA
<213> Synthetic

<400> 255
tatatatgcg gccgcgggag cccaggggat gcccag 36

<210> 256
<211> 36
<212> DNA
<213> Synthetic

<400> 256
tatatatgcg gccgccaggg gagcccaggg gatgcc 36

<210> 257
<211> 36
<212> DNA
<213> Synthetic

<400> 257
tatatatgcg gccgcgctca ggggagccca ggggat 36

<210> 258
<211> 36
<212> DNA
<213> Synthetic

<400> 258
tatatatgcg gccgcggagc tcaggggagc ccaggg 36

<210> 259
<211> 36
<212> DNA
<213> Synthetic

<400> 259
tatatatgcg gccgcgcagg agctcagggg agccca 36

<210> 260
<211> 36
<212> DNA
<213> Synthetic

<400> 260
tatatatgcg gccgcggggc aggagctcag gggagc 36

<210> 261
<211> 36
<212> DNA
<213> Synthetic

<400> 261
tatatatgcg gccgcgctgg ggcaggagct cagggg 36

<210> 262
<211> 36
<212> DNA
<213> Synthetic

<400> 262
tatatatgcg gccgcctggc tggggcagga gctcag 36

<210> 263
<211> 36
<212> DNA
<213> Synthetic

<400> 263
tatatatgcg gccgcggcct ggctggggca ggagct 36

<210> 264
<211> 36
<212> DNA
<213> Synthetic

<400> 264
tatatatgcg gccgccaggg cctggctggg gcagga 36

<210> 265
<211> 36
<212> DNA
<213> Synthetic

<400> 265
tatatatgcg gccgcctgca gggcctggct ggggca 36

<210> 266
<211> 36
<212> DNA
<213> Synthetic

<400> 266
tatatatgcg gccgccagct gcagggcctg gctggg 36

<210> 267
<211> 36
<212> DNA
<213> Synthetic

<400> 267
tatatatgcg gccgctgcca gctgcagggc ctggct 36

<210> 268
<211> 36
<212> DNA
<213> Synthetic

<400> 268
tatatatgcg gccgcgcctg ccagctgcag ggcctg 36

<210> 269
<211> 36
<212> DNA
<213> Synthetic

<400> 269
tatatatgcg gccgcgcagc ctgccagctg cagggc 36

<210> 270
<211> 36
<212> DNA
<213> Synthetic

<400> 270
tatatatgcg gccgccaagc agcctgccag ctgcag 36

<210> 271
<211> 36
<212> DNA
<213> Synthetic

<400> 271
tatatatgcg gccgcgctca agcagcctgc cagctg 36

<210> 272
<211> 36
<212> DNA
<213> Synthetic

<400> 272
tatatatgcg gccgcgttgc tcaagcagcc tgccag 36

<210> 273
<211> 36
<212> DNA
<213> Synthetic

<400> 273
tatatatgcg gccgcgaggt tgctcaagca gcctgc 36

<210> 274
<211> 36

<212> DNA
<213> Synthetic

<400> 274
tatatatgcg gccgcatgga ggttgctcaa gcagcc 36

<210> 275
<211> 36
<212> DNA
<213> Synthetic

<400> 275
tatatatgcg gccgcgctat ggaggttgct caagca 36

<210> 276
<211> 36
<212> DNA
<213> Synthetic

<400> 276
tatatatgcg gccgcgccgc tatggaggtt gctcaa 36

<210> 277
<211> 36
<212> DNA
<213> Synthetic

<400> 277
tatatatgcg gccgcaaggc cgctatggag gttgct 36

<210> 278
<211> 36
<212> DNA
<213> Synthetic

<400> 278
tatatatgcg gccgcgaaaa ggccgctatg gaggtt 36

<210> 279
<211> 36
<212> DNA
<213> Synthetic

<400> 279
tatatatgcg gccgcgagga aaaggccgct atggag 36

<210> 280
<211> 36
<212> DNA
<213> Synthetic

<400> 280
tatatatgcg gccgcgtaga ggaaaaggcc gctatg 36

<210> 281
<211> 36
<212> DNA
<213> Synthetic

<400> 281
tatatatgcg gccgcctggt agaggaaaag gccgct 36

<210> 282
<211> 36
<212> DNA
<213> Synthetic

<400> 282
tatatatgcg gccgcccccc ggtagaggaa aaggcc 36

<210> 283
<211> 36
<212> DNA
<213> Synthetic

<400> 283
tatatatgcg gccgcgagcc cctggtagag gaaaag 36

<210> 284
<211> 36
<212> DNA
<213> Synthetic

<400> 284
tatatatgcg gccgccagga gcccctggta gaggaa 36

<210> 285
<211> 36
<212> DNA
<213> Synthetic

<400> 285
tatatatgcg gccgcctgca ggagcccctg gtagag 36

<210> 286
<211> 36
<212> DNA
<213> Synthetic

<400> 286
tatatatgcg gccgcggcct gcaggagccc ctggta 36

<210> 287
<211> 36
<212> DNA
<213> Synthetic

<400> 287
tatatatgcg gccgccaggg cctgcaggag cccctg 36

<210> 288
<211> 36
<212> DNA
<213> Synthetic

<400> 288
tatatatgcg gccgcttcca gggcctgcag gagccc 36

<210> 289
<211> 36
<212> DNA
<213> Synthetic

<400> 289
tatatatgcg gccgcccctt ccagggcctg caggag 36

<210> 290
<211> 36
<212> DNA
<213> Synthetic

<400> 290
tatatatgcg gccgctatcc cttccagggc ctgcag 36

<210> 291
<211> 36
<212> DNA
<213> Synthetic

<400> 291
tatatatgcg gccgcggata tcccttccag ggcctg 36

<210> 292
<211> 36
<212> DNA
<213> Synthetic

<400> 292
tatatatgcg gccgcggggg atatcccttc cagggc 36

<210> 293
<211> 36
<212> DNA
<213> Synthetic

<400> 293
tatatatgcg gccgcctcgg gggatatccc ttccag 36

<210> 294
<211> 36
<212> DNA
<213> Synthetic

<400> 294
tatatatgcg gccgccaact cgggggatat cccttc 36

<210> 295
<211> 36
<212> DNA
<213> Synthetic

<400> 295
tatatatgcg gccgcaccca actcggggga tatccc 36

<210> 296
<211> 36
<212> DNA
<213> Synthetic

<400> 296
tatatatgcg gccgcgggac ccaactcggg ggatat 36

<210> 297
<211> 36
<212> DNA
<213> Synthetic

<400> 297
tatatatgcg gccgcggtgg gacccaactc gggggga 36

<210> 298
<211> 36
<212> DNA
<213> Synthetic

<400> 298
tatatatgcg gccgccaagg tgggacccaa ctcggg 36

<210> 299
<211> 36
<212> DNA
<213> Synthetic

<400> 299
tatatatgcg gccgcgtcca aggtgggacc caactc 36

<210> 300

<211> 36
<212> DNA
<213> Synthetic

<400> 300
tatatatgcg gccgctgtgt ccaaggtggg acccaa 36

<210> 301
<211> 36
<212> DNA
<213> Synthetic

<400> 301
tatatatgcg gccgccagtg tgtccaaggt gggacc 36

<210> 302
<211> 36
<212> DNA
<213> Synthetic

<400> 302
tatatatgcg gccgcctgca gtgtgtccaa ggtggg 36

<210> 303
<211> 36
<212> DNA
<213> Synthetic

<400> 303
tatatatgcg gccgccagct gcagtgtgtc caaggt 36

<210> 304
<211> 36
<212> DNA
<213> Synthetic

<400> 304
tatatatgcg gccgcgtcca gctgcagtgt gtccaa 36

<210> 305
<211> 36
<212> DNA
<213> Synthetic

<400> 305
tatatatgcg gccgcgacgt ccagctgcag tgtgtc 36

<210> 306
<211> 36
<212> DNA
<213> Synthetic

<400> 306
tatatatgcg gccgcggcga cgtccagctg cagtgt 36

<210> 307
<211> 36
<212> DNA
<213> Synthetic

<400> 307
tatatatgcg gccgcgtcgg cgacgtccag ctgcag 36

<210> 308
<211> 36
<212> DNA
<213> Synthetic

<400> 308

tatatatgcg gccgcaaagt cggcgacgtc cagctg 36

<210> 309
<211> 36
<212> DNA
<213> Synthetic

<400> 309
tatatatgcg gccgcggcaa agtcggcgac gtccag 36

<210> 310
<211> 36
<212> DNA
<213> Synthetic

<400> 310
tatatatgcg gccgcggtgg caaagtcggc gacgtc 36

<210> 311
<211> 36
<212> DNA
<213> Synthetic

<400> 311
tatatatgcg gccgcggtgg tggcaaagtc ggcgac 36

<210> 312
<211> 36
<212> DNA
<213> Synthetic

<400> 312
tatatatgcg gccgcgatgg tggtggcaaa gtcggc 36

<210> 313
<211> 36
<212> DNA
<213> Synthetic

<400> 313
tatatatgcg gccgcccaga tggtggtggc aaagtc 36

<210> 314
<211> 36
<212> DNA
<213> Synthetic

<400> 314
tatatatgcg gccgcctgcc agatggtggt ggcaaa 36

<210> 315
<211> 36
<212> DNA
<213> Synthetic

<400> 315
tatatatgcg gccgcctgct gccagatggt ggtggc 36

<210> 316
<211> 36
<212> DNA
<213> Synthetic

<400> 316
tatatatgcg gccgccatct gctgccagat ggtggt 36

<210> 317
<211> 36
<212> DNA

<211> 36
<212> DNA
<213> Synthetic

<400> 231
tatatatgcg gccgcacaca gcttctcctg gagcgc 36

<210> 232
<211> 36
<212> DNA
<213> Synthetic

<400> 232
tatatatgcg gccgcggcac acagcttctc ctggag 36

<210> 233
<211> 36
<212> DNA
<213> Synthetic

<400> 233
tatatatgcg gccgcggtgg cacacagctt ctcctg 36

<210> 234
<211> 36
<212> DNA
<213> Synthetic

<400> 234
tatatatgcg gccgcgtagg tggcacacag cttctc 36

<210> 235
<211> 36
<212> DNA
<213> Synthetic

<400> 235
tatatatgcg gccgccttgt aggtggcaca cagctt 36

<210> 236
<211> 36
<212> DNA
<213> Synthetic

<400> 236
tatatatgcg gccgccagct tgtaggtggc acacag 36

<210> 237
<211> 36
<212> DNA
<213> Synthetic

<400> 237
tatatatgcg gccgcgcaca gcttgtaggt ggcaca 36

<210> 238
<211> 36
<212> DNA
<213> Synthetic

<400> 238
tatatatgcg gccgcgtggc acagcttgta ggtggc 36

<210> 239
<211> 36
<212> DNA
<213> Synthetic

<400> 239

tatatatgcg gccgcggggt ggcacagctt gtaggt 36

<210> 240
<211> 36
<212> DNA
<213> Synthetic

<400> 240
tatatatgcg gccgcctcgg ggtggcacag cttgta 36

<210> 241
<211> 36
<212> DNA
<213> Synthetic

<400> 241
tatatatgcg gccgcctcct cggggtggca cagctt 36

<210> 242
<211> 36
<212> DNA
<213> Synthetic

<400> 242
tatatatgcg gccgccagct cctcggggtg gcacag 36

<210> 243
<211> 36
<212> DNA
<213> Synthetic

<400> 243
tatatatgcg gccgccacca gctcctcggg gtggca 36

<210> 244
<211> 36
<212> DNA
<213> Synthetic

<400> 244
tatatatgcg gccgccagca ccagctcctc ggggtg 36

<210> 245
<211> 36
<212> DNA
<213> Synthetic

<400> 245
tatatatgcg gccgcgagca gcaccagctc ctcggg 36

<210> 246
<211> 36
<212> DNA
<213> Synthetic

<400> 246
tatatatgcg gccgctccga gcagcaccag ctcctc 36

<210> 247
<211> 36
<212> DNA
<213> Synthetic

<400> 247
tatatatgcg gccgcgtgtc cgagcagcac cagctc 36

<210> 248
<211> 36
<212> DNA

<213> Synthetic

<400> 248
tatatatgcg gccgcagagt gtccgagcag caccag 36

<210> 249
<211> 36
<212> DNA
<213> Synthetic

<400> 249
tatatatgcg gccgccagag agtgtccgag cagcac 36

<210> 250
<211> 36
<212> DNA
<213> Synthetic

<400> 250
tatatatgcg gccgcgccca gagagtgtcc gagcag 36

<210> 251
<211> 36
<212> DNA
<213> Synthetic

<400> 251
tatatatgcg gccgcgatgc ccagagagtg tccgag 36

<210> 252
<211> 36
<212> DNA
<213> Synthetic

<400> 252
tatatatgcg gccgcgggga tgcccagaga gtgtcc 36

<210> 253
<211> 36
<212> DNA
<213> Synthetic

<400> 253
tatatatgcg gccgcccagg ggatgcccag agagtg 36

<210> 254
<211> 36
<212> DNA
<213> Synthetic

<400> 254
tatatatgcg gccgcagccc aggggatgcc cagaga 36

<210> 255
<211> 36
<212> DNA
<213> Synthetic

<400> 255
tatatatgcg gccgcgggag cccaggggat gcccag 36

<210> 256
<211> 36
<212> DNA
<213> Synthetic

<400> 256
tatatatgcg gccgccaggg gagcccaggg gatgcc 36

<210> 257
<211> 36
<212> DNA
<213> Synthetic

<400> 257
tatatatgcg gccgcgctca ggggagccca ggggat 36

<210> 258
<211> 36
<212> DNA
<213> Synthetic

<400> 258
tatatatgcg gccgcggagc tcaggggagc ccaggg 36

<210> 259
<211> 36
<212> DNA
<213> Synthetic

<400> 259
tatatatgcg gccgcgcagg agctcagggg agccca 36

<210> 260
<211> 36
<212> DNA
<213> Synthetic

<400> 260
tatatatgcg gccgcggggc aggagctcag gggagc 36

<210> 261
<211> 36
<212> DNA
<213> Synthetic

<400> 261
tatatatgcg gccgcgctgg ggcaggagct cagggg 36

<210> 262
<211> 36
<212> DNA
<213> Synthetic

<400> 262
tatatatgcg gccgcctggc tggggcagga gctcag 36

<210> 263
<211> 36
<212> DNA
<213> Synthetic

<400> 263
tatatatgcg gccgcggcct ggctggggca ggagct 36

<210> 264
<211> 36
<212> DNA
<213> Synthetic

<400> 264
tatatatgcg gccgccaggg cctggctggg gcagga 36

<210> 265
<211> 36
<212> DNA
<213> Synthetic

<400> 265
tatatatgcg gccgcctgca gggcctggct ggggca 36

<210> 266
<211> 36
<212> DNA
<213> Synthetic

<400> 266
tatatatgcg gccgccagct gcagggcctg gctggg 36

<210> 267
<211> 36
<212> DNA
<213> Synthetic

<400> 267
tatatatgcg gccgctgcca gctgcagggc ctggct 36

<210> 268
<211> 36
<212> DNA
<213> Synthetic

<400> 268
tatatatgcg gccgcgcctg ccagctgcag ggcctg 36

<210> 269
<211> 36
<212> DNA
<213> Synthetic

<400> 269
tatatatgcg gccgcgcagc ctgccagctg cagggc 36

<210> 270
<211> 36
<212> DNA
<213> Synthetic

<400> 270
tatatatgcg gccgccaagc agcctgccag ctgcag 36

<210> 271
<211> 36
<212> DNA
<213> Synthetic

<400> 271
tatatatgcg gccgcgctca agcagcctgc cagctg 36

<210> 272
<211> 36
<212> DNA
<213> Synthetic

<400> 272
tatatatgcg gccgcgttgc tcaagcagcc tgccag 36

<210> 273
<211> 36
<212> DNA
<213> Synthetic

<400> 273
tatatatgcg gccgcgaggt tgctcaagca gcctgc 36

<210> 274
<211> 36

<212> DNA
<213> Synthetic

<400> 274
tatatatgcg gccgcatgga ggttgctcaa gcagcc 36

<210> 275
<211> 36
<212> DNA
<213> Synthetic

<400> 275
tatatatgcg gccgcgctat ggaggttgct caagca 36

<210> 276
<211> 36
<212> DNA
<213> Synthetic

<400> 276
tatatatgcg gccgcgccgc tatggaggtt gctcaa 36

<210> 277
<211> 36
<212> DNA
<213> Synthetic

<400> 277
tatatatgcg gccgcaaggc cgctatggag gttgct 36

<210> 278
<211> 36
<212> DNA
<213> Synthetic

<400> 278
tatatatgcg gccgcgaaaa ggccgctatg gaggtt 36

<210> 279
<211> 36
<212> DNA
<213> Synthetic

<400> 279
tatatatgcg gccgcgagga aaaggccgct atggag 36

<210> 280
<211> 36
<212> DNA
<213> Synthetic

<400> 280
tatatatgcg gccgcgtaga ggaaaaggcc gctatg 36

<210> 281
<211> 36
<212> DNA
<213> Synthetic

<400> 281
tatatatgcg gccgcctggt agaggaaaag gccgct 36

<210> 282
<211> 36
<212> DNA
<213> Synthetic

<400> 282
tatatatgcg gccgcccccc ggtagaggaa aaggcc 36

<210> 283
<211> 36
<212> DNA
<213> Synthetic

<400> 283
tatatatgcg gccgcgagcc cctggtagag gaaaag 36

<210> 284
<211> 36
<212> DNA
<213> Synthetic

<400> 284
tatatatgcg gccgccagga gcccctggta gaggaa 36

<210> 285
<211> 36
<212> DNA
<213> Synthetic

<400> 285
tatatatgcg gccgcctgca ggagcccctg gtagag 36

<210> 286
<211> 36
<212> DNA
<213> Synthetic

<400> 286
tatatatgcg gccgcggcct gcaggagccc ctggta 36

<210> 287
<211> 36
<212> DNA
<213> Synthetic

<400> 287
tatatatgcg gccgccaggg cctgcaggag cccctg 36

<210> 288
<211> 36
<212> DNA
<213> Synthetic

<400> 288
tatatatgcg gccgcttcca gggcctgcag gagccc 36

<210> 289
<211> 36
<212> DNA
<213> Synthetic

<400> 289
tatatatgcg gccgcccctt ccagggcctg caggag 36

<210> 290
<211> 36
<212> DNA
<213> Synthetic

<400> 290
tatatatgcg gccgctatcc cttccagggc ctgcag 36

<210> 291
<211> 36
<212> DNA
<213> Synthetic

<400> 291
tatatatgcg gccgcggata tcccttccag ggcctg 36

<210> 292
<211> 36
<212> DNA
<213> Synthetic

<400> 292
tatatatgcg gccgcggggg atatcccttc cagggc 36

<210> 293
<211> 36
<212> DNA
<213> Synthetic

<400> 293
tatatatgcg gccgcctcgg gggatatccc ttccag 36

<210> 294
<211> 36
<212> DNA
<213> Synthetic

<400> 294
tatatatgcg gccgccaact cgggggatat cccttc 36

<210> 295
<211> 36
<212> DNA
<213> Synthetic

<400> 295
tatatatgcg gccgcaccca actcggggga tatccc 36

<210> 296
<211> 36
<212> DNA
<213> Synthetic

<400> 296
tatatatgcg gccgcgggac ccaactcggg ggatat 36

<210> 297
<211> 36
<212> DNA
<213> Synthetic

<400> 297
tatatatgcg gccgcggtgg gacccaactc ggggga 36

<210> 298
<211> 36
<212> DNA
<213> Synthetic

<400> 298
tatatatgcg gccgccaagg tgggacccaa ctcggg 36

<210> 299
<211> 36
<212> DNA
<213> Synthetic

<400> 299
tatatatgcg gccgcgtcca aggtgggacc caactc 36

<210> 300

<211> 36
<212> DNA
<213> Synthetic

<400> 300
tatatatgcg gccgctgtgt ccaaggtggg acccaa 36

<210> 301
<211> 36
<212> DNA
<213> Synthetic

<400> 301
tatatatgcg gccgccagtg tgtccaaggt gggacc 36

<210> 302
<211> 36
<212> DNA
<213> Synthetic

<400> 302
tatatatgcg gccgcctgca gtgtgtccaa ggtggg 36

<210> 303
<211> 36
<212> DNA
<213> Synthetic

<400> 303
tatatatgcg gccgccagct gcagtgtgtc caaggt 36

<210> 304
<211> 36
<212> DNA
<213> Synthetic

<400> 304
tatatatgcg gccgcgtcca gctgcagtgt gtccaa 36

<210> 305
<211> 36
<212> DNA
<213> Synthetic

<400> 305
tatatatgcg gccgcgacgt ccagctgcag tgtgtc 36

<210> 306
<211> 36
<212> DNA
<213> Synthetic

<400> 306
tatatatgcg gccgcggcga cgtccagctg cagtgt 36

<210> 307
<211> 36
<212> DNA
<213> Synthetic

<400> 307
tatatatgcg gccgcgtcgg cgacgtccag ctgcag 36

<210> 308
<211> 36
<212> DNA
<213> Synthetic

<400> 308

tatatatgcg gccgcaaagt cggcgacgtc cagctg 36

<210> 309
<211> 36
<212> DNA
<213> Synthetic

<400> 309
tatatatgcg gccgcggcaa agtcggcgac gtccag 36

<210> 310
<211> 36
<212> DNA
<213> Synthetic

<400> 310
tatatatgcg gccgcggtgg caaagtcggc gacgtc 36

<210> 311
<211> 36
<212> DNA
<213> Synthetic

<400> 311
tatatatgcg gccgcggtgg tggcaaagtc ggcgac 36

<210> 312
<211> 36
<212> DNA
<213> Synthetic

<400> 312
tatatatgcg gccgcgatgg tggtggcaaa gtcggc 36

<210> 313
<211> 36
<212> DNA
<213> Synthetic

<400> 313
tatatatgcg gccgcccaga tggtggtggc aaagtc 36

<210> 314
<211> 36
<212> DNA
<213> Synthetic

<400> 314
tatatatgcg gccgcctgcc agatggtggt ggcaaa 36

<210> 315
<211> 36
<212> DNA
<213> Synthetic

<400> 315
tatatatgcg gccgcctgct gccagatggt ggtggc 36

<210> 316
<211> 36
<212> DNA
<213> Synthetic

<400> 316
tatatatgcg gccgccatct gctgccagat ggtggt 36

<210> 317
<211> 36
<212> DNA

<213> Synthetic

<400> 317
tatatatgcg gccgcttcca tctgctgcca gatggt 36

<210> 318
<211> 36
<212> DNA
<213> Synthetic

<400> 318
tatatatgcg gccgcttctt ccatctgctg ccagat 36

<210> 319
<211> 36
<212> DNA
<213> Synthetic

<400> 319
tatatatgcg gccgccagtt cttccatctg ctgcca 36

<210> 320
<211> 36
<212> DNA
<213> Synthetic

<400> 320
tatatatgcg gccgctccca gttcttccat ctgctg 36

<210> 321
<211> 36
<212> DNA
<213> Synthetic

<400> 321
tatatatgcg gccgccattc ccagttcttc catctg 36

<210> 322
<211> 36
<212> DNA
<213> Synthetic

<400> 322
tatatatgcg gccgcggcca ttcccagttc ttccat 36

<210> 323
<211> 36
<212> DNA
<213> Synthetic

<400> 323
tatatatgcg gccgcagggg ccattcccag ttcttc 36

<210> 324
<211> 36
<212> DNA
<213> Synthetic

<400> 324
tatatatgcg gccgcggcag gggccattcc cagttc 36

<210> 325
<211> 36
<212> DNA
<213> Synthetic

<400> 325
tatatatgcg gccgccaggg caggggccat tcccag 36

<210> 326
<211> 36
<212> DNA
<213> Synthetic

<400> 326
tatatatgcg gccgcctgca gggcaggggc cattcc 36

<210> 327
<211> 36
<212> DNA
<213> Synthetic

<400> 327
tatatatgcg gccgcgggct gcagggcagg ggccat 36

<210> 328
<211> 36
<212> DNA
<213> Synthetic

<400> 328
tatatatgcg gccgcggtgg gctgcagggc aggggc 36

<210> 329
<211> 36
<212> DNA
<213> Synthetic

<400> 329
tatatatgcg gccgcctggg tgggctgcag ggcagg 36

<210> 330
<211> 36
<212> DNA
<213> Synthetic

<400> 330
tatatatgcg gccgcaccct gggtgggctg cagggc 36

<210> 331
<211> 36
<212> DNA
<213> Synthetic

<400> 331
tatatatgcg gccgcggcac cctgggtggg ctgcag 36

<210> 332
<211> 36
<212> DNA
<213> Synthetic

<400> 332
tatatatgcg gccgccatgg caccctgggt gggctg 36

<210> 333
<211> 36
<212> DNA
<213> Synthetic

<400> 333
tatatatgcg gccgccggca tggcaccctg ggtggg 36

<210> 334
<211> 36
<212> DNA
<213> Synthetic

<400> 334
tatatatgcg gccgcggccg gcatggcacc ctgggt 36

<210> 335
<211> 36
<212> DNA
<213> Synthetic

<400> 335
tatatatgcg gccgcgaagg ccggcatggc accctg 36

<210> 336
<211> 36
<212> DNA
<213> Synthetic

<400> 336
tatatatgcg gccgcggcga aggccggcat ggcacc 36

<210> 337
<211> 36
<212> DNA
<213> Synthetic

<400> 337
tatatatgcg gccgcagagg cgaaggccgg catggc 36

<210> 338
<211> 36
<212> DNA
<213> Synthetic

<400> 338
tatatatgcg gccgcagcag aggcgaaggc cggcat 36

<210> 339
<211> 36
<212> DNA
<213> Synthetic

<400> 339
tatatatgcg gccgcgaaag cagaggcgaa ggccgg 36

<210> 340
<211> 36
<212> DNA
<213> Synthetic

<400> 340
tatatatgcg gccgcctgga aagcagaggc gaaggc 36

<210> 341
<211> 36
<212> DNA
<213> Synthetic

<400> 341
tatatatgcg gccgcgcgct ggaaagcaga ggcgaa 36

<210> 342
<211> 36
<212> DNA
<213> Synthetic

<400> 342
tatatatgcg gccgcccggc gctggaaagc agaggc 36

<210> 343
<211> 36

<212> DNA
<213> Synthetic

<400> 343
tatatatgcg gccgctgccc ggcgctggaa agcaga 36

<210> 344
<211> 36
<212> DNA
<213> Synthetic

<400> 344
tatatatgcg gccgctcctg cccggcgctg gaaagc 36

<210> 345
<211> 36
<212> DNA
<213> Synthetic

<400> 345
tatatatgcg gccgcccctc ctgcccggcg ctggaa 36

<210> 346
<211> 36
<212> DNA
<213> Synthetic

<400> 346
tatatatgcg gccgcgaccc ctcctgcccg gcgctg 36

<210> 347
<211> 36
<212> DNA
<213> Synthetic

<400> 347
tatatatgcg gccgccagga cccctcctgc ccggcg 36

<210> 348
<211> 36
<212> DNA
<213> Synthetic

<400> 348
tatatatgcg gccgcaacca ggacccctcc tgcccg 36

<210> 349
<211> 36
<212> DNA
<213> Synthetic

<400> 349
tatatatgcg gccgcagcaa ccaggacccc tcctgc 36

<210> 350
<211> 36
<212> DNA
<213> Synthetic

<400> 350
tatatatgcg gccgcgctag caaccaggac ccctcc 36

<210> 351
<211> 36
<212> DNA
<213> Synthetic

<400> 351
tatatatgcg gccgcatggc tagcaaccag gacccc 36

<210> 352
<211> 36
<212> DNA
<213> Synthetic

<400> 352
tatatatgcg gccgccagat ggctagcaac caggac 36

<210> 353
<211> 36
<212> DNA
<213> Synthetic

<400> 353
tatatatgcg gccgcctgca gatggctagc aaccag 36

<210> 354
<211> 36
<212> DNA
<213> Synthetic

<400> 354
tatatatgcg gccgcgctct gcagatggct agcaac 36

<210> 355
<211> 36
<212> DNA
<213> Synthetic

<400> 355
tatatatgcg gccgcgaagc tctgcagatg gctagc 36

<210> 356
<211> 36
<212> DNA
<213> Synthetic

<400> 356
tatatatgcg gccgccagga agctctgcag atggct 36

<210> 357
<211> 36
<212> DNA
<213> Synthetic

<400> 357
tatatatgcg gccgcctcca ggaagctctg cagatg 36

<210> 358
<211> 36
<212> DNA
<213> Synthetic

<400> 358
tatatatgcg gccgccacct ccaggaagct ctgcag 36

<210> 359
<211> 36
<212> DNA
<213> Synthetic

<400> 359
tatatatgcg gccgccgaca cctccaggaa gctctg 36

<210> 360
<211> 36
<212> DNA
<213> Synthetic

<400> 360
tatatatgcg gccgcgtacg acacctccag gaagct 36

<210> 361
<211> 36
<212> DNA
<213> Synthetic

<400> 361
tatatatgcg gccgcgcggt acgacacctc caggaa 36

<210> 362
<211> 36
<212> DNA
<213> Synthetic

<400> 362
tatatatgcg gccgcaacgc ggtacgacac ctccag 36

<210> 363
<211> 36
<212> DNA
<213> Synthetic

<400> 363
tatatatgcg gccgctagaa cgcggtacga cacctc 36

<210> 364
<211> 36
<212> DNA
<213> Synthetic

<400> 364
tatatatgcg gccgcgcgta gaacgcggta cgacac 36

<210> 365
<211> 36
<212> DNA
<213> Synthetic

<400> 365
tatatatgcg gccgcgtggc gtagaacgcg gtacga 36

<210> 366
<211> 36
<212> DNA
<213> Synthetic

<400> 366
tatatatgcg gccgcaaggt ggcgtagaac gcggta 36

<210> 367
<211> 36
<212> DNA
<213> Synthetic

<400> 367
tatatatgcg gccgccgcaa ggtggcgtag aacgcg 36

<210> 368
<211> 39
<212> DNA
<213> Synthetic

<400> 368
tatatatgcg gccgcctgcg caaggtggcg tagaacgcg 39

INTERNATIONAL SEARCH REPORT

Intern nal Application No
PCT/US 99/20891

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/10 C12N15/25 C12N15/62 C07K14/535 C07K14/005

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MCWHERTER C (REPRINT) ET AL: "Protein engineering of the myelopoietins by mutation, circular permutation and combination of amino acid sequences." BLOOD, (15 NOV 1997) VOL. 90, NO. 10, PART 2, SUPP. '1!, PP. 3531-3531. PUBLISHER: W B SAUNDERS CO, INDEPENDENCE SQUARE WEST CURTIS CENTER, STE 300, PHILADELPHIA, PA 19106-3399. , XP000876993	14,15
Y	Abstract no. 3531 abstract	1-13

—

-/—

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
24 February 2000	14/03/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Hornig, H

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/20891

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MCKEARN J (REPRINT) ET AL: "Structure and function of myelopoietins, a family of hematopoietic growth factors." BLOOD, (15 NOV 1997) VOL. 90, NO. 10, PART 1, SUPP. '1!, PP. 249-249. PUBLISHER: W B SAUNDERS CO, INDEPENDENCE SQUARE WEST CURTIS CENTER, STE 300, PHILADELPHIA, PA 19106-3399. , XP000876981	14,15
Y	Abstract no .249 abstract	1-13
X	WO 97 12977 A (SEARLE & CO ;ZURFLUH LINDA L (US); KLEIN BARBARA K (US); MCWHERTER) 10 April 1997 (1997-04-10)	14
Y	the whole document	1-13
X	S. MERLIN ET AL.: "Phage presentation and affinity-selection of a deletion mutant of human interleukin-3" APPLIED BIOCHEMISTRY AND BIOTECHNOLOGY, vol. 67, no. 3, September 1997 (1997-09), pages 199-214, XP000884149 HUMANA PRESS INC., CLIFTON, N.J.,US	14
Y	the whole document	1-13
Y	H. GRAM ET AL.: "Phage display as a rapid gene expression system: production of bioactive cytokine-phage and generation of neutralizing monoclonal antibodies" J. IMMUNOL. METHODS, vol. 161, 1993, pages 169-176, XP002131429 ELSEVIER, AMSTERDAM NL the whole document	1-13
Y	CLACKSON T ET AL: "IN VITRO SELECTION FROM PROTEIN AND PEPTIDE LIBRARIES" TIBTECH,GB,CAMBRIDGE, vol. 12, 1 May 1994 (1994-05-01), pages 173-184, XP000652419 cited in the application the whole document	1-13
X	R. GRAF AND H.K. SCHACHMAN: "Random circular permutation of genes and expressed polypeptide chains: Application of the method to the catalytic chains of aspartate transcarbamoylase" PROC. NATL. ACAD. SCI., vol. 93, October 1996 (1996-10), pages 11591-11596, XP002131430 NATL. ACAD. SCI.,WASHINGTON,DC,US; cited in the application	14
Y	the whole document	1-13

-/—

1

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	HORLICK R A ET AL: "PERMUTEINS OF INTERLEUKIN 1 BETA-A SIMPLIFIED APPROACH FOR THE CONSTRUCTION OF PERMUTATED PROTEINS HAVING NEW TERMINI" PROTEIN ENGINEERING,GB,OXFORD UNIVERSITY PRESS, SURREY, vol. 5, no. 5, 1 January 1992 (1992-01-01), pages 427-431, XP002022097 ISSN: 0269-2139 cited in the application the whole document	14
X	KREITMAN R J ET AL: "A CIRCULARLY PERMUTED RECOMBINANT INTERLEUKIN 4 TOXIN WITH INCREASED ACTIVITY" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA,US,NATIONAL ACADEMY OF SCIENCE. WASHINGTON, vol. 91, no. 15, 1 July 1994 (1994-07-01), pages 6889-6893, XP002022099 ISSN: 0027-8424 the whole document	14
P,X	MCWHERTER C A ET AL: "Circular permutation of the granulocyte colony-stimulating factor receptor agonist domain of myelopoietin." BIOCHEMISTRY, (1999 APR 6) 38 (14) 4564-71. JOURNAL CODE: AOG. , XP002131431 the whole document	14,15
T	T.J. MACVITTIE ET AL.: "Myelopoietin, an engineered chimeric IL-3 and G-CSF receptor agonist, stimulates multilineage hematopoietic recovery in a nonhuman model of radiation-induced myelosuppression" BLOOD, vol. 95, no. 3, 1 February 2000 (2000-02-01), pages 837-845, XP002131432 SAUNDERS, DULUTH, NEW YORK , US the whole document	

1

INTERNATIONAL SEARCH REPORT

information on patent family members

Intern. nal Application No
PCT/US 99/20891

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9712977 A	10-04-1997	AU 7390096 A	28-04-1997
		CA 2234042 A	10-04-1997
		EP 0859843 A	26-08-1998

<213> Synthetic

<400> 317
tatatatgcg gccgcttcca tctgctgcca gatggt 36

<210> 318
<211> 36
<212> DNA
<213> Synthetic

<400> 318
tatatatgcg gccgcttctt ccatctgctg ccagat 36

<210> 319
<211> 36
<212> DNA
<213> Synthetic

<400> 319
tatatatgcg gccgccagtt cttccatctg ctgcca 36

<210> 320
<211> 36
<212> DNA
<213> Synthetic

<400> 320
tatatatgcg gccgctccca gttcttccat ctgctg 36

<210> 321
<211> 36
<212> DNA
<213> Synthetic

<400> 321
tatatatgcg gccgccattc ccagttcttc catctg 36

<210> 322
<211> 36
<212> DNA
<213> Synthetic

<400> 322
tatatatgcg gccgcggcca ttcccagttc ttccat 36

<210> 323
<211> 36
<212> DNA
<213> Synthetic

<400> 323
tatatatgcg gccgcagggg ccattcccag ttcttc 36

<210> 324
<211> 36
<212> DNA
<213> Synthetic

<400> 324
tatatatgcg gccgcggcag gggccattcc cagttc 36

<210> 325
<211> 36
<212> DNA
<213> Synthetic

<400> 325
tatatatgcg gccgccaggg caggggccat tcccag 36

<210> 326
<211> 36
<212> DNA
<213> Synthetic

<400> 326
tatatatgcg gccgcctgca gggcaggggc cattcc 36

<210> 327
<211> 36
<212> DNA
<213> Synthetic

<400> 327
tatatatgcg gccgcgggct gcagggcagg ggccat 36

<210> 328
<211> 36
<212> DNA
<213> Synthetic

<400> 328
tatatatgcg gccgcggtgg gctgcagggc aggggc 36

<210> 329
<211> 36
<212> DNA
<213> Synthetic

<400> 329
tatatatgcg gccgcctggg tgggctgcag ggcagg 36

<210> 330
<211> 36
<212> DNA
<213> Synthetic

<400> 330
tatatatgcg gccgcaccct gggtgggctg cagggc 36

<210> 331
<211> 36
<212> DNA
<213> Synthetic

<400> 331
tatatatgcg gccgcggcac cctgggtggg ctgcag 36

<210> 332
<211> 36
<212> DNA
<213> Synthetic

<400> 332
tatatatgcg gccgccatgg caccctgggt gggctg 36

<210> 333
<211> 36
<212> DNA
<213> Synthetic

<400> 333
tatatatgcg gccgccggca tggcaccctg ggtggg 36

<210> 334
<211> 36
<212> DNA
<213> Synthetic

<400> 334
tatatatgcg gccgcggccg gcatggcacc ctgggt 36

<210> 335
<211> 36
<212> DNA
<213> Synthetic

<400> 335
tatatatgcg gccgcgaagg ccggcatggc accctg 36

<210> 336
<211> 36
<212> DNA
<213> Synthetic

<400> 336
tatatatgcg gccgcggcga aggccggcat ggcacc 36

<210> 337
<211> 36
<212> DNA
<213> Synthetic

<400> 337
tatatatgcg gccgcagagg cgaaggccgg catggc 36

<210> 338
<211> 36
<212> DNA
<213> Synthetic

<400> 338
tatatatgcg gccgcagcag aggcgaaggc cggcat 36

<210> 339
<211> 36
<212> DNA
<213> Synthetic

<400> 339
tatatatgcg gccgcgaaag cagaggcgaa ggccgg 36

<210> 340
<211> 36
<212> DNA
<213> Synthetic

<400> 340
tatatatgcg gccgcctgga aagcagaggc gaaggc 36

<210> 341
<211> 36
<212> DNA
<213> Synthetic

<400> 341
tatatatgcg gccgcgcgct ggaaagcaga ggcgaa 36

<210> 342
<211> 36
<212> DNA
<213> Synthetic

<400> 342
tatatatgcg gccgcccggc gctggaaagc agaggc 36

<210> 343
<211> 36

<212> DNA
<213> Synthetic

<400> 343
tatatatgcg gccgctgccc ggcgctggaa agcaga 36

<210> 344
<211> 36
<212> DNA
<213> Synthetic

<400> 344
tatatatgcg gccgctcctg cccggcgctg gaaagc 36

<210> 345
<211> 36
<212> DNA
<213> Synthetic

<400> 345
tatatatgcg gccgcccctc ctgcccggcg ctggaa 36

<210> 346
<211> 36
<212> DNA
<213> Synthetic

<400> 346
tatatatgcg gccgcgaccc ctcctgcccg gcgctg 36

<210> 347
<211> 36
<212> DNA
<213> Synthetic

<400> 347
tatatatgcg gccgccagga cccctcctgc ccggcg 36

<210> 348
<211> 36
<212> DNA
<213> Synthetic

<400> 348
tatatatgcg gccgcaacca ggacccctcc tgcccg 36

<210> 349
<211> 36
<212> DNA
<213> Synthetic

<400> 349
tatatatgcg gccgcagcaa ccaggacccc tcctgc 36

<210> 350
<211> 36
<212> DNA
<213> Synthetic

<400> 350
tatatatgcg gccgcgctag caaccaggac ccctcc 36

<210> 351
<211> 36
<212> DNA
<213> Synthetic

<400> 351
tatatatgcg gccgcatggc tagcaaccag gacccc 36

<210> 352
<211> 36
<212> DNA
<213> Synthetic

<400> 352
tatatatgcg gccgccagat ggctagcaac caggac 36

<210> 353
<211> 36
<212> DNA
<213> Synthetic

<400> 353
tatatatgcg gccgcctgca gatggctagc aaccag 36

<210> 354
<211> 36
<212> DNA
<213> Synthetic

<400> 354
tatatatgcg gccgcgctct gcagatggct agcaac 36

<210> 355
<211> 36
<212> DNA
<213> Synthetic

<400> 355
tatatatgcg gccgcgaagc tctgcagatg gctagc 36

<210> 356
<211> 36
<212> DNA
<213> Synthetic

<400> 356
tatatatgcg gccgccagga agctctgcag atggct 36

<210> 357
<211> 36
<212> DNA
<213> Synthetic

<400> 357
tatatatgcg gccgcctcca ggaagctctg cagatg 36

<210> 358
<211> 36
<212> DNA
<213> Synthetic

<400> 358
tatatatgcg gccgccacct ccaggaagct ctgcag 36

<210> 359
<211> 36
<212> DNA
<213> Synthetic

<400> 359
tatatatgcg gccgccgaca cctccaggaa gctctg 36

<210> 360
<211> 36
<212> DNA
<213> Synthetic

<400> 360
tatatatgcg gccgcgtacg acacctccag gaagct 36

<210> 361
<211> 36
<212> DNA
<213> Synthetic

<400> 361
tatatatgcg gccgcgcggt acgacacctc caggaa 36

<210> 362
<211> 36
<212> DNA
<213> Synthetic

<400> 362
tatatatgcg gccgcaacgc ggtacgacac ctccag 36

<210> 363
<211> 36
<212> DNA
<213> Synthetic

<400> 363
tatatatgcg gccgctagaa cgcggtacga cacctc 36

<210> 364
<211> 36
<212> DNA
<213> Synthetic

<400> 364
tatatatgcg gccgcgcgta gaacgcggta cgacac 36

<210> 365
<211> 36
<212> DNA
<213> Synthetic

<400> 365
tatatatgcg gccgcgtggc gtagaacgcg gtacga 36

<210> 366
<211> 36
<212> DNA
<213> Synthetic

<400> 366
tatatatgcg gccgcaaggt ggcgtagaac gcggta 36

<210> 367
<211> 36
<212> DNA
<213> Synthetic

<400> 367
tatatatgcg gccgccgcaa ggtggcgtag aacgcg 36

<210> 368
<211> 39
<212> DNA
<213> Synthetic

<400> 368
tatatatgcg gccgcctgcg caaggtggcg tagaacgcg 39

INTERNATIONAL SEARCH REPORT

Intern nal Application No
PCT/US 99/20891

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/10 C12N15/25 C12N15/62 C07K14/535 C07K14/005

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MCWHERTER C (REPRINT) ET AL: "Protein engineering of the myelopoietins by mutation, circular permutation and combination of amino acid sequences." BLOOD, (15 NOV 1997) VOL. 90, NO. 10, PART 2, SUPP. '1!, PP. 3531-3531. PUBLISHER: W B SAUNDERS CO, INDEPENDENCE SQUARE WEST CURTIS CENTER, STE 300, PHILADELPHIA, PA 19106-3399. , XP000876993	14,15
Y	Abstract no. 3531 abstract	1-13

—/—

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
24 February 2000	14/03/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Hornig, H

1

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MCKEARN J (REPRINT) ET AL: "Structure and function of myelopoietins, a family of hematopoietic growth factors." BLOOD, (15 NOV 1997) VOL. 90, NO. 10, PART 1, SUPP. '1!, PP. 249-249. PUBLISHER: W B SAUNDERS CO, INDEPENDENCE SQUARE WEST CURTIS CENTER, STE 300, PHILADELPHIA, PA 19106-3399. , XP000876981	14,15
Y	Abstract no .249 abstract	1-13
X	WO 97 12977 A (SEARLE & CO ;ZURFLUH LINDA L (US); KLEIN BARBARA K (US); MCWHERTER) 10 April 1997 (1997-04-10)	14
Y	the whole document	1-13
X	S. MERLIN ET AL.: "Phage presentation and affinity-selection of a deletion mutant of human interleukin-3" APPLIED BIOCHEMISTRY AND BIOTECHNOLOGY, vol. 67, no. 3, September 1997 (1997-09), pages 199-214, XP000884149 HUMANA PRESS INC., CLIFTON, N.J.,US	14
Y	the whole document	1-13
Y	H. GRAM ET AL.: "Phage display as a rapid gene expression system: production of bioactive cytokine-phage and generation of neutralizing monoclonal antibodies" J. IMMUNOL. METHODS, vol. 161, 1993, pages 169-176, XP002131429 ELSEVIER, AMSTERDAM NL the whole document	1-13
Y	CLACKSON T ET AL: "IN VITRO SELECTION FROM PROTEIN AND PEPTIDE LIBRARIES" TIBTECH,GB,CAMBRIDGE, vol. 12, 1 May 1994 (1994-05-01), pages 173-184, XP000652419 cited in the application the whole document	1-13
X	R. GRAF AND H.K. SCHACHMAN: "Random circular permutation of genes and expressed polypeptide chains: Application of the method to the catalytic chains of aspartate transcarbamoylase" PROC. NATL. ACAD. SCI., vol. 93, October 1996 (1996-10), pages 11591-11596, XP002131430 NATL. ACAD. SCI.,WASHINGTON,DC,US; cited in the application	14
Y	the whole document	1-13

-/—

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	HORLICK R A ET AL: "PERMUTEINS OF INTERLEUKIN 1 BETA-A SIMPLIFIED APPROACH FOR THE CONSTRUCTION OF PERMUTATED PROTEINS HAVING NEW TERMINI" PROTEIN ENGINEERING,GB,OXFORD UNIVERSITY PRESS, SURREY, vol. 5, no. 5, 1 January 1992 (1992-01-01), pages 427-431, XP002022097 ISSN: 0269-2139 cited in the application the whole document	14
X	KREITMAN R J ET AL: "A CIRCULARLY PERMUTED RECOMBINANT INTERLEUKIN 4 TOXIN WITH INCREASED ACTIVITY" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA,US,NATIONAL ACADEMY OF SCIENCE. WASHINGTON, vol. 91, no. 15, 1 July 1994 (1994-07-01), pages 6889-6893, XP002022099 ISSN: 0027-8424 the whole document	14
P,X	MCWHERTER C A ET AL: "Circular permutation of the granulocyte colony-stimulating factor receptor agonist domain of myelopoietin." BIOCHEMISTRY, (1999 APR 6) 38 (14) 4564-71. JOURNAL CODE: A0G. , XP002131431 the whole document	14,15
T	T.J. MACVITTIE ET AL.: "Myelopoietin, an engineered chimeric IL-3 and G-CSF receptor agonist, stimulates multilineage hematopoietic recovery in a nonhuman model of radiation-induced myelosuppression" BLOOD, vol. 95, no. 3, 1 February 2000 (2000-02-01), pages 837-845, XP002131432 SAUNDERS, DULUTH, NEW YORK , US the whole document	

1

INTERNATIONAL SEARCH REPORT

information on patent family members

Intern. nal Application No
PCT/US 99/20891

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9712977 A	10-04-1997	AU 7390096 A	28-04-1997
		CA 2234042 A	10-04-1997
		EP 0859843 A	26-08-1998

CORRECTED VERSION*

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/10, 15/25, 15/62, C07K 14/535, 14/005 | A1

(11) International Publication Number: WO 00/18905

(43) International Publication Date: 6 April 2000 (06.04.00)

(21) International Application Number: PCT/US99/20891

(22) International Filing Date: 24 September 1999 (24.09.99)

(30) Priority Data:
60/101,908 25 September 1998 (25.09.98) US

(71) Applicant (*for all designated States except US*): G.D. SEARLE & CO. [US/US]; Corporate Patent Department, P.O. Box 5110, Chicago, IL 60680–5110 (US).

(72) Inventor; and

(75) Inventor/Applicant (*for US only*): LEE, Stephen, C. [US/US]; 828 Fernview Drive, Creve Coeur, MO 63141 (US).

(74) Agents: BAUER, S., Christopher et al.; G.D. Searle & Co., Corporate Patent Dept., P.O. Box 5110, Chicago, IL 60680–5110 (US).

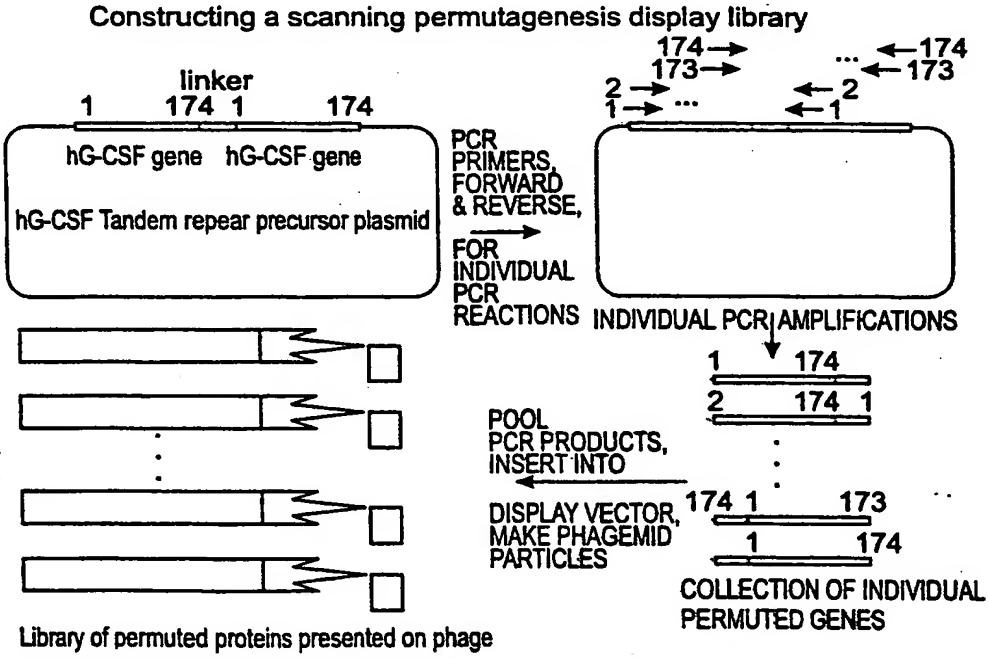
(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: METHOD OF PRODUCING PERMUTEINS BY SCANNING PERMUTAGENESIS



(57) Abstract

A method of producing circularly–permuted proteins (permuteins) by scanning permutagenesis comprises making and inserting a series of circularly–permuted genes into a display vector, expressing these genes such that the gene products are localized to the surface of the display vector, generating a library of display vectors presenting the permuted protein, affinity–selecting the display vectors with a target protein that can bind the permuted protein, isolating and analyzing clones of selected display vectors to identify the circularly–permuted protein. The invention further discloses methods of expressing and uses of permuteins.

*(Referred to in PCT Gazette No. 43/2000, Section II)

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

Method of producing permuteins by scanning permutagenesis

Priority

The present application claims priority under Title 35, United States Code, § 119 of United States Provisional Application Serial No. 60/101,908, filed September 25, 1998.

Field of the invention

A method of producing circularly-permuted proteins (permuteins) by scanning permutagenesis comprises making and inserting a series of circularly-permuted genes into a display vector, expressing these genes such that the gene products are localized to the surface of the display vector, generating a library of display vectors presenting the permuted protein, affinity-selecting the display vectors with a target protein that can bind the permuted protein, isolating and analyzing clones of selected display vectors to identify the circularly-permuted protein. The invention further discloses methods of expressing and uses of permuteins.

Background of the invention

Protein permutagenesis

Circularly permuted proteins are made by reordering the primary sequence of a parent protein. The amino and carboxy terminal ends of the parent protein are joined by a peptide linker and new amino and carboxy terminal ends are generated at other positions in the sequence. This technique of generating variants has been applied to a wide variety of proteins (Table 1).

Circularly permuted proteins, in many cases, are structurally and functionally similar to their non-permuted parent molecule after they undergo refolding. The information necessary to direct the folding of proteins into tertiary structures is present in secondary structural domains. Vectorial folding of proteins from their native amino to carboxy ends is not often observed. The ability of permuteins to retain structural and functional properties is remarkable. extending earlier observations on the plasticity of proteins with respect to amino acid

substitutions (Olins P.O. et al., *J. Biol. Chem.* 270: 23754-23760, 1995; Lowman and Wells, *J. Mol. Biol.* 234: 564-578, 1993) and short amino acid insertions (Sondek, J. and D. Shortle, *Proteins* 7: 387-393, 1990; Shortle, D. and J. Sondek, *Curr. Opin. Biotechnol.* 6: 299-305).

Protein sequence reorganization

Rearrangements of DNA sequences serve an important role in evolution by generating a diversity of new proteins differing in structure and function. Gene duplication and exon shuffling, for example, generate diversity and provide organisms with a competitive advantage since the basal mutation rate is low (Doolittle, *Protein Science* 1: 191-200, 1992).

Recombinant DNA methods have facilitated studies on the effect of sequence transposition on protein folding, structure, and function. The first rearrangement of proteins using this approach was described by Goldenberg and Creighton (*J. Mol. Biol.* 165:407-413, 1983). A new N-terminus is selected at an internal site (breakpoint) of the original sequence, the new sequence having the same order of amino acids as the original from the breakpoint until it reaches an amino acid that is at or near the original C-terminus. At this point the new sequence is joined, either directly or through an additional portion of sequence (linker), to an amino acid that is at or near the original N-terminus, and the new sequence continues with the same sequence as the original until it reaches a point that is at or near the amino acid that was N-terminal to the breakpoint site of the original sequence, this residue forming the new C-terminus of the chain. Similar approaches have also been used in other studies (Cunningham et al., *Proc. Natl. Acad. Sci. U.S.A.* 76:3218-3222, 1979; Teather & Erfle, *J. Bacteriol.* 172: 3837-3841, 1990; Schimming et al., *Eur. J. Biochem.* 204: 13-19, 1992; Yamiuchi and Minamikawa, *FEBS Lett.* 260:127-130, 1991: MacGregor et al., *FEBS Lett.* 378:263-266, 1996).

These general approaches have been applied to proteins which range in size from 58 to 462 amino acids (Goldenberg & Creighton, *J. Mol. Biol.* 165:407-413, 1983; Li & Coffino, *Mol. Cell. Biol.* 13:2377-2383, 1993). The proteins represent a broad range of structural classes, including proteins that contain predominantly alpha helix (interleukin-4; Kreitman et al., *Cytokine* 7:311-318, 1995), beta sheet (interleukin-1; Horlick et al., *Protein Eng.* 5:427-431, 1992), or mixtures of the two types of secondary structures (yeast phosphoribosyl anthranilate isomerase; Luger et al., *Science* 243:206-210, 1989).

Although broad categories of protein function are represented in these sequence reorganization studies, the results of these studies have been highly variable. In many cases substantially lower activity, solubility, or thermodynamic stability were observed (*E. coli* dihydrofolate reductase, aspartate transcarbamoylase, phosphoribosyl anthranilate isomerase, glyceraldehyde-3-phosphate dehydrogenase, ornithine decarboxylase, ompA, yeast phosphoglycerate dehydrogenase). In other cases, the sequence rearranged protein appeared to have many nearly identical properties as its natural counterpart (basic pancreatic trypsin inhibitor, T4 lysozyme, ribonuclease T1, Bacillus β-glucanase, interleukin-1β, α-spectrin SH3 domain, pepsinogen, interleukin-4). In exceptional cases, an unexpected improvement over some properties of the natural sequence was observed, e.g., the solubility and refolding rate for rearranged α-spectrin SH3 domain sequences, and the receptor affinity and anti-tumor activity of transposed interleukin-4-*Pseudomonas* exotoxin fusion molecule (Kreitman et al., *Proc. Natl. Acad. Sci. U.S.A.* 91:6889-6893, 1994; Kreitman et al., *Cancer Res.* 55:3357-3363, 1995).

The primary motivation for reorganization studies has been to study the role of short-range and long-range interactions in protein folding and stability. Sequence rearrangements of this type convert a subset of interactions that are long-range in the original sequence into short-range interactions in the new sequence, and vice versa. The fact that many of these sequence rearrangements are able to attain a conformation with at least some activity is persuasive evidence that protein folding occurs by multiple folding pathways (Viguera et al., *J. Mol. Biol.* 247:670-681, 1995). In the case of the SH3 domain of alpha-spectrin, choosing new termini at locations that corresponded to beta hairpin turns resulted in proteins with slightly less stability, but which were nevertheless able to fold.

The positions of the internal breakpoints used in the studies cited above are found exclusively on the surface of proteins, and are distributed throughout the linear sequence without any obvious bias towards the ends or the middle (the variation in the relative distance from the original N-terminus to the breakpoint is ca. 10 to 80% of the total sequence length). The linkers connecting the original N- and C-termini in these studies have ranged from 0 to 9 residues. In one case (Yang & Schachman, *Proc. Natl. Acad. Sci. U.S.A.* 90:11980-11984, 1993), a portion of sequence has been deleted from the original C-terminal segment, and the connection made from the truncated C-terminus to the original N-terminus. Flexible hydrophilic residues such as Gly and Ser are frequently used in the linkers. Viguera et al.(*J. Mol. Biol.* 247:670-681, 1995) compared joining the

original N- and C- termini with 3- or 4-residue linkers; the 3-residue linker was less thermodynamically stable. Protasova et al. (*Protein Eng.* 7:1373-1377, 1994) used 3- or 5-residue linkers in connecting the original N-termini of *E. coli* dihydrofolate reductase; only the 3-residue linker produced protein in good yield.

Protein permutagenesis can be used to optimize the activity of fusion proteins or proteins conjugated to other molecules. A fusion between interleukin-4 (IL-4) and *Pseudomonas* exotoxin has been permuted resulting in a protein that has the first amino acid of the IL-4 domain at position 38 and the new carboxy end occurs at amino acid position 37 (Kreitman, R. J. et al., *Proc Natl Acad Sci USA* 91: 6889-6893, 1994). The permuted fusion has increased affinity for the IL-4 receptor, increased cytotoxicity to IL-4 receptor bearing renal carcinoma cells, and increased anti-tumor activity in a murine model, compared to the non-permuted parent fusion protein (Kreitman, R. J. et al., *Proc Natl Acad Sci USA* 91: 6889-6893, 1994; Kreitman, R. J. et al., *Cancer Res.* 55:3357-3363, 1995; Puri, R. K. et al., *Cellular Immunol.* 171: 80-86, 1996). Increased potency of the permuted molecule is believed to result from a reduction in steric interference between the IL-4 domain in the parent molecule and its receptor.

Steric hindrance is likely to be a concern for other chimeric proteins which interact with receptors through a relatively large area of their surface. The same issue also arises with bioconjugates, containing relatively small chemicals conjugated to proteins or other molecules in complex polymers (Rose, K. et al., *Molecular Immunology* 32: 1031-1037, 1995).

Phage display methods

Display methods allow affinity selection of protein variants from a library of displayed proteins or peptides (Clackson, T. and J.A. Wells, *Tibtech* 12: 173-184, 1994; Winter, G., *Drug Development Res.* 33: 71-89, 1994). Many biological entities can be used in display methodologies (so-called "genetic packages" for presentation, including bacterial and eukaryotic cells, various eukaryotic and prokaryotic viruses, and spores), but the most commonly used vehicles used for display are filamentous bacteriophage, as used herein. We envision the possibility that a genetic package other than the particular phage used here could be used to present libraries of permuteins, and if so, constitute essentially the same invention.

Foreign proteins are presented on the surface of a phage particle, and the gene encoding the foreign protein is encapsulated in the virion. Because they are linked by the phage particle, affinity isolation of the presented protein also leads to affinity isolation of the corresponding genes. Extremely large libraries of phage presented proteins are constructed and affinity screened very rapidly. From the standpoint of how quickly mutant proteins can be made and screened for activity, phage display is the most efficient mutagenesis technique currently available.

Functional properties of permuteins

Permuteins can have improved biological properties by acting through several mechanisms. The permutein acting on the same type of cell as its parent molecule, may have increased binding, or other action, by virtue of increased avidity. Dimers or higher order multimers of these proteins with themselves or other chemical groups, including proteins, can have increased efficacy or potency, or both.

Permuteins can also have improved therapeutic properties through a variety of mechanisms such as: (1) alterations in the overall on- or off-rates or K_a or K_d of the ligand(s) on the target cell; (2) activation or blockade of complementary receptor signaling pathways; and/or (3) more specific targeting of to the cell of interest. The permuteins may also possess a unique pharmacokinetic distribution and clearance profile (Dehmer et al., *Circulation*, 91, 2188-2194, 1995; Tanaka et al.,. *Nature Medicine*, 3, 437-442, 1997).

Permuteins can also have improved properties *in vivo,* compared to the two components individually, as a result of alterations in biodistribution or half-life. The improved properties can also result from the binding of the permutein to one or more of the receptors, pharmacokinetics, or uptake of the permutein is altered in a favorable manner.

Molecular biology approaches have traditionally been used to permute proteins (Horlick, R.A. et al., *Protein Engineering* 5: 427-431, 1992) although chemical approaches have been used to make small permuted proteins (Goldenberg, D. P. and T. E. Creighton, *J. Mol. Biol.* 165: 407-413, 1983). These approaches are relatively labor intensive, limiting the number of permuteins that can be generated and efficiently screened for the desired biological activities. Rapid methods of generating permuteins, coupled with efficient methods for screening are needed that will result in the identification of novel active molecules.

Summary of the invention

The present invention is an improved method for generating permuteins (scanning permutagenesis) based on the display of proteins on bacteriophage surface proteins. Phage display is a powerful, yet convenient tool, traditionally used for mutagenesis and screening (Clackson, T. and J.A. Wells, *Tibtech* 12: 173-184, 1994). Improvements to this technology allow the rapid generation and screening of libraries of permuteins. Variables, such as position of the new termini and the length and composition of peptide linkers can easily be varied to generate libraries of the desired diversity.

The present invention relates to methods of producing biologically-active circularly permuted proteins of the formula C^1-L^1-N^1, derived from a parent protein of the formula N^1-C^1, wherein C^1 is comprised of a segment derived from the carboxy portion of said parent protein; N^1 is comprised of a segment derived from the amino terminal portion of said parent protein; and L^1 is a chemical bond or a linker, linking C^1 to the amino terminus of L^1 and carboxy terminus of L^1 to the amino terminus of N^1; comprising the steps of: (a) making a series of circularly-permuted genes; (b) inserting said circularly-permuted genes into a display vector; (c) expressing said circularly-permuted genes such that the proteins encoded by said genes are presented on the surface of the display vector; (d) generating a library of display vectors presenting the expressed circularly permuted protein; (e) affinity-select the presenting display vectors with a target protein that can bind a biologically-active circularly-permuted protein; (f) isolate and analyze clones of selected display vectors to identify the presented circularly-permuted protein.

Preferably the method of making a series of circularly-permuted genes is selected from the group consisting of making a tandemly-repeated intermediate, total synthesis of a synthetic gene, assembly of a gene from synthetic oligonucleotides, DNA amplification, and limited digestion of a circular intermediate.

Preferably, the display vector is selected from the group consisting of bacteriophage display vectors, bacteria, and baculovirus vectors. Even more preferably the presentation vector is a bacteriophage. Even more preferably, the presentation vector is bacteriophage M13. Most preferably, the presentation vector is a bacteriophage M13 gene III vector.

Preferably the method of making a series of circularly permuted genes is a method of making a tandem repeat intermediate. Even more preferably circularly permuted genes are amplified from the repeat by gene amplification.

Preferably the method of affinity selection comprises the steps consisting of (a) binding said presentation display vectors to a target protein; (b) eluting said display vectors; (c) amplifying said display vectors; and (d) biopanning a pool of said amplified display vectors.

Preferably, the length of C^1 in the permutein is longer than the length of C^1 in said parent protein. More preferably, the length of C^1 in the permutein is shorter than the length of C^1 in said parent protein. Most preferably, the length of C^1 in the permutein is the same length as the length of C^1 in said parent protein.

Preferably, the length of N^1 in the permutein is longer than the length of N^1 in said parent protein. More preferably, the length of N^1 in the permutein is shorter than the length of N^1 in said parent protein. Most preferably, the length of N^1 in the permutein is the same length as the length of N^1 in said parent protein.

The invention also contemplates circularly permuted proteins of the formula C^1-L^1-N^1 made by the method of scanning permutagenesis. Preferably, the DNA sequence encoding said linker L^1 is selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 368.

Preferably, the circularly-permuted protein is the G-CSF receptor agonist domain of a species of mylepoietin (MPO). MPO is one member of a family of novel dual cytokine receptor agonists (McKearn, J.P., Myelorestorative activities of synthokine and myelopoietin. *In* Proceedings of the 1996 IBC Conference on Therapeutic Applications of Cytokines, pp. 4.3.1-4.3.18, 1996) which are amenable to manipulation by phage display (Merlin, S. et al., *Applied Biochemistry and Biotechnology* 67: 15-29, 1997; Lee, S.C., Phage presentation for cytokine engineering. IBC's Second International Conference on Display Technologies, 1997).

Brief description of the figures

Figure 1. Schematic depiction of scanning permutagenesis

Plate A of Figure 1 shows the strategy to generate a scanning permutagenesis phage display library. A plasmid containing directly-repeated

tandem copies of the hG-CSF gene, for example, is constructed by standard methods. The tandem repeat plasmid is used as the template for PCR amplification of genes encoding permuted proteins. Each copy of the G-CSF gene is indicated in light gray (turquoise), and a DNA segment encoding a peptide linker is indicated in dark gray (red).

In individual PCR reactions, oligonucleotide primers that initiate PCR polymerization at the first nucleotide of a chosen codon of G-CSF, and directing polymerization to the end of the tandem construct specifying the carboxy end of the protein encoded on the template is annealed to the tandem template. Also, a second specific primer is also annealed to the template that initiates polymerization at the last nucleotide of the codon encoding the amino acid immediately preceding the codon where polymerization begins with first primer, and which directs polymerization in the opposite direction from that first primer. Amplification between these two primers produces a DNA segment encoding a permuted protein. For example, amplification between the primer indicated by a black arrow initiating at codon 2 and the primer indicated by the blue arrow and initiating at the codon before 2 (codon 1) produces an amplified gene encoding a permuted protein whose amino terminal residue is amino acid 2 of the native protein, and whose final amino acid is amino acid 1 of the native protein.

A linker peptide is present between the first and final amino acids of the parent protein (residues 1 and 174 in this example). A total of 174 individual amplifications would produce a complete collection of all permuted proteins of this example. More limited collections containing only a selected set of permuteins can be made, as well as more extensive collections made from multiple tandem template plasmids, each containing a different linker sequence between the first and last residues of the two directly repeated tandem gene sequences. The collection of amplified segments can then be inserted into a phagemid presentation vector by standard methods. Phagemid particles produced from these presentation constructs are the scanning permutagenesis phage display library.

Plate B of Figure 1 shows the affinity screening of a phage display library (See Clackson, T. and J.A. Wells, *Tibtech* 12: 173-184, 1994; Winter, G., *Drug Development Res.* 33: 71-89, 1994). In this example, a hG-CSF scanning permutagenesis library as described in Figure 1A is screened using the hG-CSF receptor expressed on mammalian cells as the affinity reagent. In Figure 1B, individual presented proteins are indicated by the shaded circles or diamonds and the affinity reagent is indicated by the light gray (pink) rectangles. Presentation

library particles are exposed to affinity reagent, unbound particles are washed away, and receptor-bound particles are eluted. The eluted particles are amplified in *E. coli*, and the screening cycle is repeated. During any round of the screening cycle, the genes encoded (in the present example encoding permuted proteins) by the selected particles can be expressed and evaluated.

Figure 2. Permuteins presented in the scanning permutagenesis library

Human G-CSF (ser17) protein is depicted as a string of circles, each circle corresponding to a single amino acid residue. Amino and carboxy ends of the protein are indicated. The amino acids of helical regions are indicated by medium gray balls, while the amino acids of inter-helical loops are indicated in light gray balls (See Hill et al., *Proc. Natl. Acad. Sci. USA* 90: 5167-5171, 1993). Amino-ends of the permuteins made for presentation in the library are indicated in dark gray. Asterisks indicate the breakpoints of the presented permuteins which were isolated by affinity screening with cells expressing hG-CSF receptor as illustrated in 1B.

Figure 3. Bioactivity of permuteins identified by affinity screening of the scanning permutagenesis library

Individual permuteins were expressed transiently in mammalian cells. Permeation molecules in the culture supernatants were quantitated by ELISA, and the proliferative activity of clones was determined using BAF-3-cells dependent on G-CSF for growth. The horizontal axis indicate concentration of protein and the vertical axis indicate incorporation of tritiated thymidine.

Definitions

The following is a list of abbreviations and the corresponding meanings as used interchangeably herein:

g = gram(s)
mg = milligram(s)
ml and mL = milliliter(s)
RT = room temperature
ug and µg = microgram(s)
uL and µl = microliter(s)

The following is a list of definitions of various terms used herein:

The term "permutein" means a circularly-permuted protein: a protein in which the amino and carboxy ends of the parent protein are joined together by a peptide linker sequence of zero or more amino acids. The amino and carboxy ends of the permuted protein occur at amino acids within the parental sequence.

The terms "chemical ligation" and "conjugation" mean a chemical reaction which covalently links two similar or dissimilar functional groups together intramolecularly or intermolecularly.

The term "peptide linker' means a compound which forms a carboxamide bond between two groups having one or more peptide linkages (CONH-) and serves as a connector for the propose of amelioration of the distance or space orientation between two molecules.

The term "native sequence" refers to an amino acid or nucleic acid sequence which is identical to a wild-type or native form of a gene or protein.

The terms "mutant amino acid sequence," "mutant protein", "variant protein", "mutein", or "mutant polypeptide" refer to a polypeptide having an amino acid sequence which varies from a native sequence due to amino acid additions, deletions, substitutions, or all three, or is encoded by a nucleotide sequence from an intentionally-made variant derived from a native sequence.

Detailed description of the invention

Determination of the amino and carboxyl termini of permuteins

The present invention encompasses circularly permuted-proteins of the formula C^1-L^1-N^1 prepared by phage display techniques. The polypeptide can be joined either directly or through a linker segment. The term "directly" defines permuteins in which the polypeptide ends are joined without a linker. Thus L^1 represents a chemical bond or a linker, preferably a polypeptide segment to which both C^1 and N^1 are joined, wherein C^1 is comprised of a segment derived from the carboxy portion of the parent protein and N^1 is comprised of a segment derived from the amino terminal portion of a parent protein represented by the general formula N^1-C^1. Preferably, N^1 and C^1 in the permuted protein C^1-L^1-N^1 are the same length as in the parent protein N^1-C^1, but each may be independently shorter or longer depending on the desired structural characteristics of the permutein. Most commonly L^1 is a linear peptide in which C^1 and N^1 are joined by amide

bonds, linking C^1 to the amino terminus of L^1 and carboxy terminus of L^1 to the amino terminus of N^1.

Additional peptide sequences may also be added to facilitate purification or identification of permuteins (e.g., poly-His). A highly antigenic peptide may also be added that would enable rapid assay and facile purification of the permuteins by a specific monoclonal antibody.

Determination of the linker

The linking group (L^1) is generally a polypeptide of between 1 and 500 amino acids in length. The linkers joining the two molecules are preferably designed to (1) allow the two molecules to fold and act independently of each other, (2) not have a propensity for developing an ordered secondary structure which could interfere with the functional domains of the two proteins, (3) have minimal hydrophobic characteristics which could interact with the functional protein domains and (4) provide steric separation of C^1 and N^1 such that C^1 and N^1 could interact simultaneously with their corresponding receptors on a single cell. Typically surface amino acids in flexible protein regions include Gly, Asn and Ser. Virtually any permutation of amino acid sequences containing Gly, Asn and Ser would be expected to satisfy the above criteria for a linker sequence. Other neutral amino acids, such as Thr and Ala, may also be used in the linker sequence. Additional amino acids may also be included in the linkers due to the addition of unique restriction sites in the linker sequence to facilitate construction of the multi-functional proteins.

Preferred L^1 linkers of the present invention include sequences selected from the group of formulas:

(SEQ ID NO: 1) through SEQ ID NO: 268)

Other linkers are also contemplated by the invention. The present invention is, however, not limited by the form, size or number of linker sequences employed. The only requirement of the linker is that it does not functionally interfere with the folding and function of the individual molecules of the multi-functional protein.

Utility of permuteins

Permuteins of the present invention may exhibit useful properties such as having similar or greater biological activity when compared to a single factor or by

having improved half-life or decreased adverse side effects, or a combination of these properties.

Permuteins which have little or no activity maybe useful as antigens for the production of antibodies for use in immunology or immunotherapy, as probes or as intermediates used to construct other useful permuteins.

The permuteins of the present invention may have an improved therapeutic profile as compared to their parent molecules. For example, some permuteins of the present invention may have a similar or more potent activities relative to other compounds or proteins without having a similar or corresponding increase in side-effects. This is particularly true of multifunctional or fusion protein therapeutics, where permutation may relieve steric and other hindrances that impair the activity of the parent fusion molecules (see Kreitman, R. J. et al., *Proc Natl Acad Sci USA* **91**: 6889-6893, 1994; Kreitman, R. J. et al., *Cancer Res.* **55**:3357-3363, 1995, for examples).

A general utility of permuteins is in the area of nanoscale devices described alternatively as "nanobiological" or "nanobiotechnological." These are nanoscale devices containing both precise structure nanomaterials and biological functional components (such as proteins). Nanodevices have been the subject of several reviews (Lee, S.C., *Trends in Biotechnology*, **16**: 239-240, 1998).

Nanobiological/nanobiotechnological devices generally contain proteins covalently coupled to polymers or other non-biological precise structure materials. Issues of steric and other interferences with protein activity are applicable to proteins in nanobiological/nanobiotechnological devices and are highly analogous to the issues with multifunctional/fusion proteins discussed above. Protein permutation is fully expected to offer a viable approach to deal with these considerations, just as it does in the case of fusion proteins (Kreitman, R. J. et al., *Proc Natl Acad Sci USA* **91**: 6889-6893, 1994; Kreitman, R. J. et al.,., *Cancer Res.* **55**:3357-3363, 1995).

Examples

The following examples will illustrate the invention in greater detail, although it will be understood that the invention is not limited to these specific examples. Various other examples will be apparent to the person skilled in the art after reading the present disclosure without departing from the spirit and scope of

the invention. It is intended that all such other examples be included within the scope of the appended claims.

General Materials and Methods

General methods of cloning, expressing, and characterizing proteins are found in T. Maniatis et al., *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, 1982, and references cited therein, incorporated herein by reference; and in J. Sambrook et al., *Molecular Cloning, A Laboratory Manual*, 2nd edition, Cold Spring Harbor Laboratory, 1989, and references cited therein, incorporated herein by reference.

Unless noted otherwise, all specialty chemicals were obtained from Sigma Chemical Company (St. Louis, MO). Restriction endonucleases and T4 DNA ligase were obtained from New England Biolabs (Beverly, MA) or Boehringer Mannheim (Indianapolis, IN).

Strains, plasmids, and bacteriophage

The bacterial strains used in these studies are listed in Table 1. Plasmids and bacteriophage used or constructed in this study are listed in Tables 2 and 3, respectively.

Phage and phagemid stocks were made and manipulated as described (Kay, B.K., Winter, J., and McCafferty, J., Phage Display of Peptides and Proteins, Academic Press, San Diego, California, 1996; Merlin, S. et al., *Applied Biochemistry and Biotechnology* 67: 15-29, 1997).

Transformation of *E. coli* strains

E. coli strains (Table 1), such as DH5α™ (Life Technologies, Gaithersburg, MD) and TG1 (Amersham Corp., Arlington Heights, IL) are used for transformation of ligation reactions and are the hosts used to prepare plasmid DNA for transfecting mammalian cells. *E. coli* strains, such as JM101 (Yanisch-Perron et al., *Gene*, 33: 103-119, 1985) and MON105 (Obukowicz et al., *Appl. and Envir. Micr.*, 58: 1511-1523, 1992) can be used for expressing the multi-functional proteins of the present invention in the cytoplasm or periplasmic space.

DH5α™ Subcloning efficiency cells are purchased as competent cells and are ready for transformation using the manufacturer's protocol, while both *E. coli* strains TG1 and MON105 are rendered competent to take up DNA using a $CaCl_2$

method. Typically, 20 to 50 mL of cells are grown in LB medium (1% Bacto-tryptone, 0.5% Bacto-yeast extract, 150 mM NaCl) to a density of approximately 1.0 optical density unit at 600 nanometers (OD600) as measured by a Baush & Lomb Spectronic spectrophotometer (Rochester, NY). The cells are collected by centrifugation and resuspended in one-fifth culture volume of $CaCl_2$ solution (50 mM $CaCl_2$, 10 mM Tris-Cl, pH7.4) and are held at 4°C for 30 minutes. The cells are again collected by centrifugation and resuspended in one-tenth culture volume of $CaCl_2$ solution. Ligated DNA is added to 0.2 mL of these cells, and the samples are held at 4°C for 30-60 minutes. The samples are shifted to 42°C for two minutes and 1.0 mL of LB is added prior to shaking the samples at 37°C for one hour. Cells from these samples are spread on plates (LB medium plus 1.5% Bacto-agar) containing either ampicillin (100 micrograms/mL, ug/mL) when selecting for ampicillin-resistant transformants, or spectinomycin (75 ug/mL) when selecting for spectinomycin-resistant transformants. The plates are incubated overnight at 37°C.

Colonies are picked and inoculated into LB plus appropriate antibiotic (100 ug/mL ampicillin or 75 ug/mL spectinomycin) and are grown at 37°C while shaking.

DNA isolation and characterization

DNA constructs were made and propagated in *E. coli* using standard molecular biology techniques (Sambrook, J. et al., *Molecular Cloning, A Laboratory Manual*, 2nd edition, Cold Spring Harbor Laboratory, 1989).

Plasmid DNA can be isolated by a number of different methods and using commercially available kits known to those skilled in the art. Plasmid DNA is isolated using the Promega Wizard™ Miniprep kit (Madison, WI), the Qiagen QIAwell Plasmid isolation kits (Chatsworth, CA) or Qiagen Plasmid Midi or Mini kit. These kits follow the same general procedure for plasmid DNA isolation. Briefly, cells are pelleted by centrifugation (5000 x g), the plasmid DNA released with sequential NaOH/acid treatment, and cellular debris is removed by centrifugation (10000 x g). The supernatant (containing the plasmid DNA) is loaded onto a column containing a DNA-binding resin, the column is washed, and plasmid DNA eluted. After screening for the colonies with the plasmid of interest, the *E. coli* cells are inoculated into 50-100 ml of LB plus appropriate antibiotic for overnight growth at 37°C in an air incubator while shaking. The purified plasmid DNA is used for DNA sequencing, further restriction enzyme digestion, additional

subcloning of DNA fragments and transfection into *E. coli*, mammalian cells, or other cell types.

Sequence confirmation

DNA sequence analysis was performed using the Genesis 2000 DNA analysis system using standard methods (Prober et al., *Science* 238: 336-341, 1987).

Purified plasmid DNA is resuspended in dH_2O and its concentration is determined by measuring the absorbance at 260/280 nm in a Bausch and Lomb Spectronic 601 UV spectrometer. DNA samples are sequenced using ABI PRISM™ DyeDeoxy™ terminator sequencing chemistry (Applied Biosystems Division of Perkin Elmer Corporation, Lincoln City, CA) kits (Part Number 401388 or 402078) according to the manufacturer's suggested protocol usually modified by the addition of 5% DMSO to the sequencing mixture. Sequencing reactions are performed in a DNA thermal cycler (Perkin Elmer Corporation, Norwalk, CT) following the recommended amplification conditions. Samples are purified to remove excess dye terminators with Centri-Sep™ spin columns (Princeton Separations, Adelphia, NJ) and lyophilized. Fluorescent dye labeled sequencing reactions are resuspended in deionized formamide, and sequenced on denaturing 4.75% polyacrylamide-8M urea gels using ABI Model 373A and Model 377 automated DNA sequencers. Overlapping DNA sequence fragments are analyzed and assembled into master DNA contigs using Sequencher DNA analysis software (Gene Codes Corporation, Ann Arbor, MI).

Expression of permuted proteins in mammalian cells

To obtain sufficient protein for analysis of the bioactivity of individual MPO molecules containing cphG-CSFs, DNA segments containing individual affinity-selected MPO: cphGCSFs were subcloned into a mammalian expression vector, and expressed transiently in BHK cells as described below.

The BHK-21 cell line can be obtained from the ATCC (Rockville, MD). The cells are cultured in Dulbecco's modified Eagle media (DMEM/high-glucose), supplemented to 2 mM (mM) L-glutamine and 10% fetal bovine serum (FBS). This formulation is designated BHK growth media. Selective media is BHK growth media supplemented with 453 units/mL hygromycin B (CalBiochem, San Diego, CA). The BHK-21 cell line was previously stably transfected with the HSV transactivating protein VP16, which transactivates the IE110 promoter found on

the plasmid pMON3359 and pMON3633 and the IE175 promoter found in the plasmid pMON3360B (Hippenmeyer, P.J. and Pegg, L.E., *Curr. Opin. Biotechnol.* 6: 548-552, 1995). The VP16 protein drives expression of genes inserted behind the IE110 or IE175 promoter. BHK-21 cells expressing the transactivating protein VP16 are designated BHK-VP16. The plasmid pMON1118 expresses the hygromycin resistance gene from the SV40 promoter (Highkin et al., *Poultry Sci.*, 70: 970-981, 1991). A similar plasmid, pSV2-hph, is available from ATCC.

BHK-VP16 cells are seeded into a 60 millimeter (mm) tissue culture dish at 3 x 10^5 cells per dish 24 hours prior to transfection. Cells are transfected for 16 hours in 3 mL of "OPTIMEM"™ (Gibco-BRL, Gaithersburg, MD) containing 10 ug of plasmid DNA containing the gene of interest, 3 ug hygromycin resistance plasmid, pMON1118, and 80 ug of Gibco-BRL "LIPOFECTAMINE"™ per dish. The media is subsequently aspirated and replaced with 3 mL of growth media. At 48 hours post-transfection, media from each dish is collected and assayed for activity (transient conditioned media). The cells are removed from the dish by trypsin-EDTA, diluted 1:10, and transferred to 100 mm tissue culture dishes containing 10 mL of selective media. After approximately 7 days in selective media, resistant cells grow into colonies several millimeters in diameter. The colonies are removed from the dish with filter paper (cut to approximately the same size as the colonies and soaked in trypsin/EDTA) and transferred to individual wells of a 24 well plate containing 1 mL of selective media. After the clones are grown to confluence, the conditioned media is re-assayed, and positive clones are expanded into growth media.

Affinity selection and screening of phagemids

Affinity reagent used for the identification of functional MPO molecules containing cphG-CSF (MPO: cphG-CSF) species from the library were BHK cells expressing the hG-CSF receptor on their surface. The library pool was subjected to iterative affinity selection (four rounds) against BHK cells expressing the h-GCSF receptor using previously described techniques (Merlin, S. et al., *Applied Biochemistry and Biotechnology* 67: 15-29, 1997). Between rounds of selection, phage eluted from the affinity reagent were amplified in *E. coli* (Kay, B.K. J. Winter, and J. McCofferty, Phage Display of Peptides and Proteins, Academic Press, San Diego, California, 1996).

Expression of proteins in *E. coli*

When large-scale quantities of recombinant protein are desirable for structure-function studies, DNA segments containing individual affinity-selected MPO:cphGCSFs are subcloned into any of a variety of bacterial plasmid expression vectors, and expressed as a cytoplasmic product or as a secreted protein in *E. coli*.

E. coli strain MON105 or JM101 harboring the plasmid of interest are grown at 37°C in M9 plus casamino acids medium with shaking in an air incubator Model G25 from New Brunswick Scientific (Edison, NJ). Growth is monitored at OD_{600} until it reaches a value of 1.0 at which time nalidixic acid (10 mg/mL) in 0.1 N NaOH is added to a final concentration of 50 µg/mL, for cultures containing plasmids with the *E. coli recA* promoter driving expression of the recombinant gene. IPTG is used in place of nalidixic acid, as a chemical inducer to facilitate expression from plasmids containing the lac promoter or hybrid lac promoters. The cultures are then shaken at 37°C for three to four additional hours. A high degree of aeration is maintained throughout the culture period in order to achieve maximal production of the desired gene product. The cells are examined under a light microscope for the presence of inclusion bodies (IB). One mL aliquots of the culture are removed for analysis of protein content by boiling the pelleted cells, treating them with reducing buffer and electrophoresis via SDS-PAGE (see Maniatis et al., "Molecular Cloning: A Laboratory Manual", 1982). The culture is centrifuged (5000 x g) to pellet the cells.

Isolation of Inclusion Bodies

The cell pellet from a 330 mL *E. coli* culture is resuspended in 15 mL of sonication buffer (10 mM 2-amino-2-(hydroxymethyl) 1,3-propanediol hydrochloride (Tris-HCl), pH 8.0 + 1 mM ethylenediaminetetraacetic acid (EDTA). These resuspended cells are sonicated using the microtip probe of a Sonicator Cell Disruptor (Model W-375, Heat Systems-Ultrasonics, Inc., Farmingdale, New York). Three rounds of sonication in sonication buffer followed by centrifugation are employed to disrupt the cells and wash the inclusion bodies (IB). The first round of sonication is a 3 minute burst followed by a 1 minute burst, and the final two rounds of sonication are for 1 minute each.

Purification

The folded proteins can be affinity-purified using affinity reagents such as monoclonal antibodies or receptor subunits attached to a suitable matrix.

Purification can also be accomplished using any of a variety of chromatographic methods such as: ion exchange, gel filtration or hydrophobic chromatography or reversed phase HPLC. These and other protein purification methods are described in detail (Methods in Enzymology, Volume 182 "Guide to Protein Purification" edited by Murray Deutscher, Academic Press, San Diego, California, 1990).

Protein Characterization

The purified protein is analyzed by RP-HPLC, electrospray mass spectrometry, and SDS-PAGE. The protein quantitation is done by amino acid composition, RP-HPLC, and Bradford protein determination. In some cases tryptic peptide mapping is performed in conjunction with electrospray mass spectrometry to confirm the identity of the protein.

Baf-3/G-CSF receptor assay

Briefly, the mouse lymphoid cell line Baf3 was transfected with human granulocyte colony stimulating factor receptor (hG-CSFR) cDNA. Stable clones of Baf3 which expressed the G-CSFR and proliferated in the presence of hG-CSF were isolated and used to investigate the activity of human G-CSF receptor agonists without the influence of other human cytokine receptor responses.

The cDNA encoding hG-CSFR (a gift from Dr. Daniel C. Link (Washington University, St. Louis, MO) was released from the plasmid pEMCV.Sralpha as a *Hin*dIII/*Eco*RI (5' to 3') fragment, gel-purified, and inserted into the mammalian cell expression plasmid pcDNA3 (Invitrogen, San Diego, CA). This plasmid contains enhancer-promoter sequences from the immediate early gene of the human cytomegalovirus (CMV), a bovine growth hormone polyadenylation signal and transcription termination sequences, a neomycin resistance gene is present for the selection of G418 stable cell clones, and an ampicillin resistance gene for selection in *E. coli*. Ligation mixtures were transformed into *E. coli* strain TG1 [delta (*lac-pro*), *sup*E, *thi*, *hsd*D5/F'(*tra*D36, *pro*A$^+$B$^+$, lacI^q, *lac*ZdeltaM15] and plasmid DNA was purified using a Qiagen Midiprep Plasmid Kit. The structure of plasmid DNAs containing hG-CSFR were confirmed by restriction enzyme analysis and by automated DNA sequence analysis using an ABI sequencing machine. One of several plasmids with the correct structure was selected and given the designation pMON30298.

Baf3 cells, maintained in complete growth medium (RPMI 1640 supplemented to 10% FBS and 10% Wehi 3B supernatant as a source for mouse IL-

3), were seeded at a subconfluent cell density of 10^5 cells/ml in growth media (RPMI 1640 5% FBS; 2 mM L-glutamine) the day prior to the electroporation. The cells were collected and rinsed twice in 10 ml serum-free RPMI 1640. The cells were diluted to 10^6/ml in serum-free RPMI and 1 ml was placed into each electroporation chamber (Gibco/BRL #1608AJ). 50 ug of plasmid DNA was added to each chamber and the chambers were incubated on ice for 30 minutes prior to electroporation. The cells were electroporated on ice at a capacitance of 800 uF, 400V, fast charge, and low ohms in a BRL CellPorator. The cells were immediately removed from the chambers and placed into 10 cm dishes containing 10 ml of growth medium. The cells were allowed to recover for 48 hr in growth media prior to selection.

After the recovery period, the cells were pelleted at 1000 rpm for 10 minutes, and resuspended into 10 ml of selection medium (growth medium containing 800 ug/ml G418 sulfate (Gibco/BRL). The cells were kept in selection media, being passaged twice weekly, until only a few viable cells could be seen in the mock transfected control cell dishes (approximately 2 weeks). After an additional 2 weeks in selection media, the cells which had been electroporated with the hG-CSFR cDNA had grown to a cell density which allowed them to be tested for proliferation in the presence of hG-CSF (Fukunaga, R. et al., *EMBO J.* 10 (10): 2855-2865, 1991).

The cell proliferation assay conditions are as follows: Briefly, 25,000 cells were plated in a microtiter 96 well plate with or without cytokine in IMDM medium supplemented with BSA (50 ug/ml), human transferrin (100 ug/ml), lipid (50 ug/ml) 2-mercaptoethanol (50 uM final concentration). Each well was incubated with 0.5 uCi of ^{3}H-thymidine (16 hours) and the incorporated radioactivity was measured. Triplicate wells containing Baf3 cells were set up with 4 nM hG-CSF, 4 nM mIL-3 or media only control. Samples of different permuted proteins were tested in each assay.

Example 1: Construction of a permutein library without a linker region

Figure 1 shows a schematic of scanning permutagenesis. A plasmid construct comprising a tandem repeat of the modified human granulocyte colony stimulating factor (hG-CSF with a serine for amino acid 17) gene joined by a sequence (GCCGG, termed a zero order linker) was generated and subcloned into the plasmid pACYC177 (Chang, A.C.Y. and S.N. Cohen, *J Bacteriol.* 134: 1141-1156, 1978) using standard molecular biology methods (Sambrook, J. et al.,

Molecular Cloning, A Laboratory Manual, 2nd edition, Cold Spring Harbor Press, New York, 1989). The resultant plasmid construct (pMON15978) was linearized by restriction digestion (*Sma*I) and used as a template for PCR amplification of circularly permuted hG-CSF (cphG-CSF) genes, following the method of Horlich (Horlick, R.A. et al., *Protein Engineering* 5: 427-431, 1992. For purposes of this demonstration of the scanning permutagenesis technique, we chose to make a limited permutein library rather than one containing every possible cphG-CSF. Figure 2 shows the position of the new amino termini for each new cphG-CSF.

Individual cphG-CSF genes were inserted into phagemid presentation vector pCANTAB 5E (Pharmacia Biotech,) such that they were expressed as a part of a MPO species (Feng, Y., N. R. Staten, C. M. Baum, N. L. Summers, M. Caparon, S. C. Bauer, L. Zurfluh, J. P. McKearn, B. K. Klein, S. C. Lee, C. A. McWherter. 1997. Multi-functional hematopoeitic receptor agonists. World Patent Application WO 97/12985) which was in turn fused to the amino end of the phage geneIII product. The presented fusion protein contained, starting from its amino terminus, a hIL-3 receptor agonist, cphG-CSF, and the phage gene III product. The juncture between the presented protein and the gene III product was as previously described (Merlin, S. et al., *Applied Biochemistry and Biotechnology* 67: 15-29, 1997).

After confirmation of the structure of each phagemid construct, phagemid particles were produced for each individual cphG-CSF-presenting species (Merlin et al., 1997). Some of these lots of particles were used to individually define the affinity properties of specific presented cphG-CSF species in analytical biopanning experiments (Caparon, M. H. et al., *Molecular Diversity* 1: 241-246, 1996; Merlin, S. et al., *Applied Biochemistry and Biotechnology* 67: 15-29, 1997), but all of the phage particle lots were titered and equivalent numbers of transducing units of each particle preparation were pooled together to form the scanning permutagenesis library for hG-CSF in an MPO background. Figure 2 shows the MPO: cp hG-CSF species present in the library.

MPO : cphG-CSF 38/37 is an example of the nomenclature used to specify the identity of individual permuted proteins. It describes a MPO molecule containing a circularly permuted human G-CSF module (with the serine 17 substitution). The first amino acid of the cphG-CSF domain is amino acid 38 of the parent protein, and the last amino acid is residue 37 of the parent.

Example 2: Presentation and Affinity screening of the MPO: cphGCSF

library

MPO: cphG-CSF 38/37, is a full hG-CSF receptor agonist (McKearn, J.P., Myelorestorative activities of synthokine and myelopoietin. *In* Proceedings of the 1996 IBC Conference on Therapeutic Applications of Cytokines, pp. 4.3.1-4.3.18, 1996). It was presented on filamentous phage as a positive control to demonstrate that permuted proteins can be presented on the surface of phage particles and affinity selected. After phagemid particles were produced from this construct, they were subjected to analytical biopanning using cells expressing the hG-CSF receptor as affinity reagent.

Table 1 shows that phage presented MPO: cphG-CSF 38/37 was affinity selected by cells expressing the hG-CSF. MPO: cphGCSF 38/37-GPIII fusion was expressed, secreted and assembled into phagemid particles, and could be affinity selected by the hG-CSF receptor. Permutagenesis of a protein does not appear to impair its successful presentation.

Relative to typical phage display libraries, the complexities of cp libraries are low, containing perhaps hundreds to thousands of individuals. The demonstration library here contained about 50 distinct clones, as opposed to more typical phage libraries containing more than 10^5 individuals (reviewed in Clackson, T. and J.A. Wells, *Tibtech* **12**: 173-184, 1994).

37 randomly chosen selectants from round 1, and fewer from subsequent rounds (17, 11 and 14 were picked from rounds 2, 3 and 4, respectively) were chosen for sequence analysis. The identity of the MPO: cphG-CSFs identified in each round is shown in Figure 2.

A total of 14 MPO: cphGCSF species were identified from the output of affinity selection (Figure 2). Most of the MPO: cphGCSF species identified from the library had new carboxy and amino termini in loop segments (9 of 14 permuteins identified), rather than in clearly defined secondary structures (See Hill et al., 1993 for the hG-CSF structure). Five selectants had termini within helical domains of hG-CSF (MPO: cphG-CSFs 13/12, 19/18, 71/70, 123/122 and 159/158). For three of these molecules (MPO: cphG-CSFs 13/12, 71/70 and 123/122) their new ends lie at the outermost ends of helices, and therefore perturbation of secondary structure caused by these permuteins may be minimal. However, MPO: cphG-CSF 19/18 and MPO: cphG-CSF 159/158 have new termini well within helix 1 and helix 4 of hG-CSF, respectively.

These data parallel the observations of Graf and Schachman, who developed a limited DNase I digestion method for "random" permutagenesis (Graf, R. and H. K. Schachman, *Proc Natl Acad Sci USA* 93:11591-11596, 1996). They identified two permutein species of aspartate transcarbamoylase that introduced new amino and carboxy ends into secondary structural domains and that retained biological activity. In their work, the majority of permuteins introducing ends into secondary structures (5/7 identified) were significantly diminished in activity. In contrast, we found a several permuteins that introduced helical breaks retained activity (See Below). The method used by Graf and Schechman frequently introduces point mutations, small insertions and deletions into the permuted proteins, potentially complicating the analysis of the effects of permutagenesis.

Example 3: Biological activity of MPO: cphG-CSFs selected from the cp phage library

To obtain sufficient protein for analysis of the bioactivity of individual MPO molecules containing cphG-CSFs, DNA segments containing individual affinity-selected MPO: cphGCSFs were subcloned into a mammalian expression vector, and expressed transiently in BHK cells as described above.

The MPO: cphG-CSFs isolated from biopanning were all expressed transiently in mammalian cells and the amount of MPO: cphG-CSF in each supernatant was determined by sandwich hIl-3 ELISA (Olins P.O. et al., *J. Biol. Chem.* 270: 23754-23760, 1995). The quantitated supernatants were then assayed for G-CSF receptor agonist activity in a Baf-3/G-CSF receptor assay (Figure 3, Table 4).

All but one of the transiently expressed MPO: cphG-CSF proteins exhibited G-CSF activity equivalent to or slightly better than that of the parent MPO molecule, including those MPO: cphG-CSFs with new carboxy and amino ends within helixes. The permutein encoded by pMON16021 with a breakpoint between positions 48 and 49 did not exhibit activity in the G-CSF-dependent proliferation assay. These data suggest that most of the proteins isolated from the library are competent to bind the hG-CSF receptor and produce a proliferation signal.

All references, patents, or applications cited herein are incorporated by reference in their entirety, as if written herein.

References

Buchwalder, A. et al., *Biochemistry* **31**:1621-1630, 1992.

Caparon, M. H. et al., *Molecular Diversity* **1**: 241-246, 1996.

Chang, A.C.Y. and S.N. Cohen, *J Bacteriol.* **134**: 1141-1156, 1978.

Clackson, T. and J.A. Wells, *Tibtech* **12**: 173-184, 1994.

Feng et al., *J. Mol. Biol.* **259**: 524-551, 1996.

Feng, Y., N. R. Staten, C. M. Baum, N. L. Summers, M. Caparon, S. C. Bauer, L. Zurfluh, J. P. McKearn, B. K. Klein, S. C. Lee, C. A. McWherter. 1997. Multi-functional hematopoeitic receptor agonists. World Patent Application WO 97/12985.

Fukunaga, R., Ishizaka-Ikeda, E., Pan, C. X., Seto, Y., and Nagata, S. Functional Domains of the Granulocyte Colony-Stimulating Factor Receptor. *EMBO J.* **10** (10):2855-2865, 1991.

Goldenberg, D. P. and T. E. Creighton, *J. Mol. Biol.* **165**: 407-413, 1983.

Graf, R. and H. K. Schachman, *Proc Natl Acad Sci USA* **93**:11591-11596, 1996.

Hahn, M. et al., *Proc Natl Acad Sci USA* **91**: 10417-10421, 1994.

Highkin et al., *Poultry Sci.*, **70**: 970-981, 1991.

Hill et al., *Proc. Natl. Acad. Sci. USA* **90**: 5167-5171, 1993.

Hippenmeyer, P.J. and L.E. Pegg, *Curr. Opin. Biotechnol.* **6**: 548-552, 1995.

Horlick, R.A. et al., *Protein Engineering* **5**: 427-431, 1992.

Jelinski, L.W. Biologically related aspects of nanoparticles, nanostructured materials and nanodevices. *In* "WTEC workshop on Global Assessment of R &D Status and Trends in Nanoparticles, Nanostructured Materials and Nanodevices", International Technology Research Institute, Loyola College, Baltimore, MD., S. C., 1998.

Johnson. J. and F. M. Raushel, *Biochemistry* **35**: 10223-10233, 1996.

Kay, B.K. J. Winter, and J. McCofferty, Phage Display of Peptides and Proteins, Academic Press, San Diego, California, 1996.

Koebnik, R. and L. Kramer, *J. Mol. Biol.* **250**: 617-626, 1995.

Kreitman, R. J. et al. *Cytokine* **7**(4): 311-318, 1995.

Kreitman, R. J. et al., *Proc Natl Acad Sci USA* **91**: 6889-6893, 1994.

Kreitman, R. J. et al., *Cancer Res.* **55**:3357-3363, 1995.

Lee, S.C. Biotechnology for Nanotechnology. *Trends in Biotechnology*, **16**: 239-240, 1998.

Lee, S.C., Phage presentation for cytokine engineering. IBC's Second International Conference on Display Technologies, 1997.

Lin, X. et al., *Protein Science* **4**: 159-166, 1995.

Lowman and Wells, *J. Mol. Biol.* **234**: 564-578, 1993.

Luger et al., *Protein Engineering* **3**: 249-258, 1990.

Luger et al., *Science* **243**: 206-210, 1989.

Maniatis, T. et al., *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, 1982.

McKearn, J.P. Myelorestorative activities of synthokine and myelopoietin. In Proceedings of the 1996 IBC Conference on Therapeutic Applications of Cytokines, pp. 4.3.1-4.3.18, 1996.

Merlin, S. et al., *Applied Biochemistry and Biotechnology* **67**: 15-29, 1997.

Mullins, L. S. et al., *J Am. Chem Soc.* **116**: 5529-5533, 1994.

Murray Deutscher (ed), Methods in Enzymology, Volume 182 "Guide to Protein Purification," Academic Press, San Diego, California, 1990.

Obukowicz, M. et al., *Appl. and Envir. Micr.*, 58: 1511-1523, 1992.

Olins P.O. et al., *J. Biol. Chem.* **270**: 23754-23760, 1995.

Prober, J.M. et al., *Science* **238**: 336-341, 1987.

Protosova, N. Y. et al., *Protein Engineering* 7: 1373-1377, 1994.

Puri, R. K. et al., *Cellular Immunol.* 171: 80-86, 1996.

Rose, K. et al., *Molecular Immunology* 32: 1031-1037, 1995.

Sambrook, J. et al., *Molecular Cloning, A Laboratory Manual,* 2nd edition, Cold Spring Harbor Press, New York, 1989.

Shortle, D. and J. Sondek, *Curr. Opin. Biotechnol.* 6: 299-305.

Smith, G. P. , *Curr. Opin. in Biotechnol.* 2: 668-673, 1991.

Smith, G. P., *Science* 228: 1315-1317, 1985.

Sondek, J. and D. Shortle, *Proteins* 7: 387-393, 1990.

Thomas, J. W. et al., *Proc Natl Acad Sci USA* 92: 3779-3783, 1995.

Winter, G., *Drug Development Res.* 33: 71-89, 1994.

Yang, Y. R. and H. K. Schachman, *Proc Natl Acad Sci USA* 90: 11980-11984, 1993.

Yanisch-Perron et al., *Gene*, 33: 103-119, 1985.

Zhang, T. et al., *Biochemistry* 32: 12311-12318, 1993.

Tables

Table 1: Circularly permuted proteins

Protein	Reference
Enzymes	
T4 lysozyme	Zhang et al., *Biochemistry* **32**:12311-12318 (1993); Zhang et al., *Nature Struct. Biol.* **1**:434-438 (1995)
dihydrofolate reductase	Buchwalder et al., *Biochemistry* **31**:1621-1630 (1994); Protasova et al., *Prot. Eng.* **7**:1373-1377 (1995)
ribonuclease T1	Mullins et al., *J. Am. Chem. Soc.* **116**:5529-5533 (1994); Garrett et al., *Protein Science* **5**:204-211 (1996)
Bacillus β-glucanase	Hahn et al., *Proc. Natl. Acad. Sci. U.S.A.* **91**:10417-10421 (1994)
aspartate transcarbamoylase	Yang and Schachman, *Proc. Natl. Acad.* Transcarbamoylase *Sci. U.S.A.* **90**:11980-11984 (1993)
phosphoribosyl anthranilate isomerase	Luger et al., *Science* **243**:206-210 (1989); Luger et al., *Prot. Eng.* **3**:249-258 (1990)
pepsin/pepsinogen	Lin et al., *Protein Science* **4**:159-166 (1995)
glyceraldehyde-3-phosphate dehydrogenase	Vignais et al., *Protein Science* **4**:994-1000 (1995)
ornithine decarboxylase	Li & Coffino, *Mol. Cell. Biol.* **13**:2377-2383 (1993)
yeast phosphoglycerate dehydrogenase	Ritco-Vonsovici et al., *Biochemistry* **34**:16543-16551 (1995)
Enzyme Inhibitor	
basic pancreatic trypsin inhibitor	Goldenberg & Creighton, *J. Mol. Biol.* **165**:407-413 (1983)
Cytokines	
interleukin-1β	Horlick et al., *Protein Eng.* **5**:427-431 (1992)
interleukin-4	Kreitman et al., *Cytokine* **7**:311-318 (1995)
Tyrosine Kinase Recognition Domain	

α-spectrin SH3 domain	Viguera et al., *J. Mol. Biol.* 247:670-681 (1995)
Transmembrane Protein	
omp A	Koebnik & Krämer, *J. Mol. Biol.* 250:617-626 (1995)
Chimeric Protein	
interleukin-4-*Pseudomonas* exotoxin fusion molecule	Kreitman et al., *Proc. Natl. Acad. Sci. U.S.A.* 91:6889-6893 (1994)

Table 2: Strains

Designation	Description or Genotype	Reference/Source
DH5α™	F⁻, *phi*80 d*lac*ZdeltaM15, delta(*lac*ZYA-*arg*F)U169, *deo*R, *rec*A1, *end*A1, *hsd*R17 (rk⁻,mk⁺), *pho*A, *sup*E44, lambda-, *thi*-1, *gyr*A96, relA1	Life Technologies, Rockville, Maryland
JM101 (ATCC# 33876)	delta (*pro lac*), *sup*E, *thi*, F'(*tra*D36, *pro*A⁺B⁺, *lac*I^q, *lac*ZdeltaM15)	Yanisch-Perron et al., *Gene*, 33: 103-119, 1985
MON105 (ATCC# 55204)	F⁻, lambda-,IN (rrnD, rrnE)1, rpoD⁺, *rpo*H358	Obukowicz et al., *Appl. and Envir. Micr.*, 58: 1511-1523, 1992
MON208	W3110 *rpo*H358, *lac*I^q, *omp*T::*kan*	Alan Easton
TG1	delta(lac-pro), *sup*E, *thi*-1, *hsd*D5/F'(*tra*D36, *pro*A⁺B⁺, *lac*I^q, *lac*ZdeltaM15)	Amersham Corp., Arlington Heights, Illinois
W3110	IN (*rrn*D-*rrn*E)1, *rph*1	Lab collection

Table 3: Plasmids

Plasmid	SEQ ID NO.	Selectable Marker	Description	Source
pACYC177		Kan^R Amp^R	Plasmid with multiple cloning sites and two selectable markers	Chang, A.C.Y. and S.N. Cohen, *J Bacteriol.* 134: 1141-1156, 1978
pMON15978		Amp^R	Plasmid construct comprising a tandem repeat of the modified human granulocyte colony stimulating factor (hG-CSF with a serine for amino acid 17) gene joined by a sequence (GCCGG, termed a zero order linker), subcloned into the plasmid pACYC177	This work
pCANTAB 5E		Amp^R	Phage display vector containing lac promoter operably linked to fd gene 3 signal sequence, a linker region, an E tag, and an fd gene 3 structural gene all cloned into the vector backbone of pUC119 containing ColE1 ori, the beta lactamase resistance gene, and an M13 ori.	Pharmacia Biotech, Piscataway, NJ
pMON16016		Amp^R	Phagemid presentation vector pCANTAB 5E derivation containing inserted individual cphG-CSF gene such that it was expressed as a part of an	This work

		MPO species, fused in turn to the amino terminus end of the phage geneIII product. The first amino acid of the cphG-CSF domain is amino acid 1 of the parent, and the last amino acid is residue 174 of the parent. The zero order linker is attached at the carboxyl end of amino acid 174.	
pMON16017	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 3 of the parent, and the last amino acid is residue 2 of the parent.	This work
pMON16029	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 7 of the parent, and the last amino acid is residue 6 of the parent.	This work
pMON16030	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 9 of the parent, and the last amino acid is residue 8 of the parent.	This work
pMON16018	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 11 of the parent, and the last amino acid is residue 10 of the parent.	This work

pMON16019	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 13 of the parent, and the last amino acid is residue 12 of the parent.	This work
pMON16031	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 15 of the parent, and the last amino acid is residue 14 of the parent.	This work
pMON16020	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 19 of the parent, and the last amino acid is residue 18 of the parent.	This work
pMON16032	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 22 of the parent, and the last amino acid is residue 21 of the parent.	This work
pMON16033	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 27 of the parent, and the last amino acid is residue 26 of the parent.	This work
pMON16034	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 31 of the parent, and the	This work

		last amino acid is residue 30 of the parent.	
pMON16035	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 35 of the parent, and the last amino acid is residue 34 of the parent.	This work
pMON16036	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 37 of the parent, and the last amino acid is residue 36 of the parent.	This work
pMON16037	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 38 of the parent, and the last amino acid is residue 37 of the parent.	This work
pMON16038	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 39 of the parent, and the last amino acid is residue 38 of the parent.	This work
pMON16039	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 43 of the parent, and the last amino acid is residue 42 of the parent.	This work
pMON16040	Amp^R	Identical to pMON16016 except the first amino acid of the	This work

		cphG-CSF domain is amino acid 45 of the parent, and the last amino acid is residue 44 of the parent.	
pMON16041	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 47 of the parent, and the last amino acid is residue 46 of the parent.	This work
pMON16022	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 49 of the parent, and the last amino acid is residue 48 of the parent.	This work
pMON16042	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 51 of the parent, and the last amino acid is residue 50 of the parent.	This work
pMON16043	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 53 of the parent, and the last amino acid is residue 52 of the parent.	This work
pMON16044	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 56 of the parent, and the last amino acid is residue 55 of the parent.	This work

pMON16023	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 60 of the parent, and the last amino acid is residue 59 of the parent.	This work
pMON16045	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 64 of the parent, and the last amino acid is residue 63 of the parent.	This work
pMON16024	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 67 of the parent, and the last amino acid is residue 66 of the parent.	This work
pMON16046	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 69 of the parent, and the last amino acid is residue 68 of the parent.	This work
pMON16025	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 71 of the parent, and the last amino acid is residue 70 of the parent.	This work
pMON16047	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 73 of the parent, and the	This work

		last amino acid is residue 72 of the parent.	
pMON16048	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 84 of the parent, and the last amino acid is residue 83 of the parent.	This work
pMON16049	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 98 of the parent, and the last amino acid is residue 97 of the parent.	This work
pMON16050	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 100 of the parent, and the last amino acid is residue 99 of the parent.	This work
pMON16051	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 102 of the parent, and the last amino acid is residue 101 of the parent.	This work
pMON16052	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 112 of the parent, and the last amino acid is residue 111 of the parent.	This work
pMON16053	Amp^R	Identical to pMON16016 except the first amino acid of the	This work

		cphG-CSF domain is amino acid 121 of the parent, and the last amino acid is residue 120 of the parent.	
pMON16026	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 123 of the parent, and the last amino acid is residue 122 of the parent.	This work
pMON16027	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 125 of the parent, and the last amino acid is residue 124 of the parent.	This work
pMON16054	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 133 of the parent, and the last amino acid is residue 132 of the parent.	This work
pMON16055	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 142 of the parent, and the last amino acid is residue 141 of the parent.	This work
pMON16056	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 143 of the parent, and the last amino acid is residue 142 of the parent.	This work

pMON16057	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 147 of the parent, and the last amino acid is residue 146 of the parent.	This work
pMON16028	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 159 of the parent, and the last amino acid is residue 158 of the parent.	This work
pMON16058	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 168 of the parent, and the last amino acid is residue 167 of the parent.	This work
pMON16059	Amp^R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 170 of the parent, and the last amino acid is residue 169 of the parent.	This work

Table 4: Analytical biopanning

Before receptor*	After receptor*	Enrichment
$1/6.6x10^{4}$	$1/6.5x10^{-1}$	990-fold

* Amp^{R}/Kan^{R} resistant colonies

Analytical biopanning shows that MPO molecules containing permuted hG-CSF domains can be presented and affinity selected in a hG-CSF receptor dependent fashion. A mixture of phagemids presenting MPO: cphG-CSF 38/37 (ampicillin resistant) and M13k07 (kanamycin resistant) were exposed to BHK cells with or without the hG-CSF receptor on their surface, washed and eluted from the cell surface. Eluted phage were introduced into *E. coli* and the transfected cells were plated on media containing kanamycin or ampicillin. The ratio of ampicillin resistant to kanamycin resistant particles were determined prior to and following exposure to receptor by counting resistant colonies.

Table 5: Activity of selected permuteins

Plasmid	Permutein breakpoint in G-CSF amino acid sequence	Activity in G-CSF-dependent proliferation assay
pMON16017	3/2	+
pMON16018	11/10	+
pMON16019	13/12	+
pMON16020	19/18	+
pMON16021	49/48	–
pMON16022	60/59	+
pMON16023	67/66	+
pMON16024	69/68	+
pMON16025	71/70	+
pMON16026	123/122	+
pMON16027	125/124	+
pMON16028	159/158	+

Table 6: SEQ ID Number/SEQ ID Name Correlation

SEQ ID NO.	SEQ ID Name	Sequence									
1.	FGS1	CGCGCGC	ACATG	TCT	ACA	CCA	TTG	GGC	CCT	GCC	AGC
2.	FGS2	CGCGCGC	ACATG	TCT	CCA	TTG	GGC	CCT	GCC	AGC	TCC
3.	FGS3	CGCGCGC	ACATG	TCT	TTG	GGC	CCT	GCC	AGC	TCC	CTG
4.	FGS4	CGCGCGC	ACATG	TCT	GGC	CCT	GCC	AGC	TCC	CTG	CCC
5.	FGS5	CGCGCGC	ACATG	TCT	CCT	GCC	AGC	TCC	CTG	CCC	CAG
6.	FGS6	CGCGCGC	ACATG	TCT	GCC	AGC	TCC	CTG	CCC	CAG	AGC
7.	FGS7	CGCGCGC	ACATG	TCT	AGC	TCC	CTG	CCC	CAG	AGC	TTC
8.	FGS8	CGCGCGC	ACATG	TCT	TCC	CTG	CCC	CAG	AGC	TTC	CTG
9.	FGS9	CGCGCGC	ACATG	TCT	CTG	CCC	CAG	AGC	TTC	CTG	CTC
10.	FGS10	CGCGCGC	ACATG	TCT	CCC	CAG	AGC	TTC	CTG	CTC	AAG
11.	FGS11	CGCGCGC	ACATG	TCT	CAG	AGC	TTC	CTG	CTC	AAG	TCT
12.	FGS12	CGCGCGC	ACATG	TCT	AGC	TTC	CTG	CTC	AAG	TCT	TTA
13.	FGS13	CGCGCGC	ACATG	TCT	TTC	CTG	CTC	AAG	TCT	TTA	GAG
14.	FGS14	CGCGCGC	ACATG	TCT	CTG	CTC	AAG	TCT	TTA	GAG	CAA
15.	FGS15	CGCGCGC	ACATG	TCT	CTC	AAG	TCT	TTA	GAG	CAA	GTG
16.	FGS16	CGCGCGC	ACATG	TCT	AAG	TCT	TTA	GAG	CAA	GTG	AGG
17.	FGS17	CGCGCGC	ACATG	TCT	TCT	TTA	GAG	CAA	GTG	AGG	AAG
18.	FGS18	CGCGCGC	ACATG	TCT	TTA	GAG	CAA	GTG	AGG	AAG	ATC
19.	FGS19	CGCGCGC	ACATG	TCT	GAG	CAA	GTG	AGG	AAG	ATC	CAG
20.	FGS20	CGCGCGC	ACATG	TCT	CAA	GTG	AGG	AAG	ATC	CAG	GGC
21.	FGS21	CGCGCGC	ACATG	TCT	GTG	AGG	AAG	ATC	CAG	GGC	GAT
22.	FGS22	CGCGCGC	ACATG	TCT	AGG	AAG	ATC	CAG	GGC	GAT	GGC
23.	FGS23	CGCGCGC	ACATG	TCT	AAG	ATC	CAG	GGC	GAT	GGC	GCA
24.	FGS24	CGCGCGC	ACATG	TCT	ATC	CAG	GGC	GAT	GGC	GCA	GCG
25.	FGS25	CGCGCGC	ACATG	TCT	CAG	GGC	GAT	GGC	GCA	GCG	CTC
26.	FGS26	CGCGCGC	ACATG	TCT	GGC	GAT	GGC	GCA	GCG	CTC	CAG
27.	FGS27	CGCGCGC	ACATG	TCT	GAT	GGC	GCA	GCG	CTC	CAG	GAG
28.	FGS28	CGCGCGC	ACATG	TCT	GGC	GCA	GCG	CTC	CAG	GAG	AAG
29.	FGS29	CGCGCGC	ACATG	TCT	GCA	GCG	CTC	CAG	GAG	AAG	CTG
30.	FGS30	CGCGCGC	ACATG	TCT	GCG	CTC	CAG	GAG	AAG	CTG	TGT
31.	FGS31	CGCGCGC	ACATG	TCT	CTC	CAG	GAG	AAG	CTG	TGT	GCC
32.	FGS32	CGCGCGC	ACATG	TCT	CAG	GAG	AAG	CTG	TGT	GCC	ACC
33.	FGS33	CGCGCGC	ACATG	TCT	GAG	AAG	CTG	TGT	GCC	ACC	TAC
34.	FGS34	CGCGCGC	ACATG	TCT	AAG	CTG	TGT	GCC	ACC	TAC	AAG
35.	FGS35	CGCGCGC	ACATG	TCT	CTG	TGT	GCC	ACC	TAC	AAG	CTG
36.	FGS36	CGCGCGC	ACATG	TCT	TGT	GCC	ACC	TAC	AAG	CTG	TGC
37.	FGS37	CGCGCGC	ACATG	TCT	GCC	ACC	TAC	AAG	CTG	TGC	CAC
38.	FGS38	CGCGCGC	ACATG	TCT	ACC	TAC	AAG	CTG	TGC	CAC	CCC
39.	FGS39	CGCGCGC	ACATG	TCT	TAC	AAG	CTG	TGC	CAC	CCC	GAG
40.	FGS40	CGCGCGC	ACATG	TCT	AAG	CTG	TGC	CAC	CCC	GAG	GAG
41.	FGS41	CGCGCGC	ACATG	TCT	CTG	TGC	CAC	CCC	GAG	GAG	CTG
42.	FGS42	CGCGCGC	ACATG	TCT	TGC	CAC	CCC	GAG	GAG	CTG	GTG
43.	FGS43	CGCGCGC	ACATG	TCT	CAC	CCC	GAG	GAG	CTG	GTG	CTG
44.	FGS44	CGCGCGC	ACATG	TCT	CCC	GAG	GAG	CTG	GTG	CTG	CTC
45.	FGS45	CGCGCGC	ACATG	TCT	GAG	GAG	CTG	GTG	CTG	CTC	GGA
46.	FGS46	CGCGCGC	ACATG	TCT	GAG	CTG	GTG	CTG	CTC	GGA	CAC
47.	FGS47	CGCGCGC	ACATG	TCT	CTG	GTG	CTG	CTC	GGA	CAC	TCT
48.	FGS48	CGCGCGC	ACATG	TCT	GTG	CTG	CTC	GGA	CAC	TCT	CTG
49.	FGS49	CGCGCGC	ACATG	TCT	CTG	CTC	GGA	CAC	TCT	CTG	GGC
50.	FGS50	CGCGCGC	ACATG	TCT	CTC	GGA	CAC	TCT	CTG	GGC	ATC
51.	FGS51	CGCGCGC	ACATG	TCT	GGA	CAC	TCT	CTG	GGC	ATC	CCC
52.	FGS52	CGCGCGC	ACATG	TCT	CAC	TCT	CTG	GGC	ATC	CCC	TGG
53.	FGS53	CGCGCGC	ACATG	TCT	TCT	CTG	GGC	ATC	CCC	TGG	GCT
54.	FGS54	CGCGCGC	ACATG	TCT	CTG	GGC	ATC	CCC	TGG	GCT	CCC
55.	FGS55	CGCGCGC	ACATG	TCT	GGC	ATC	CCC	TGG	GCT	CCC	CTG
56.	FGS56	CGCGCGC	ACATG	TCT	ATC	CCC	TGG	GCT	CCC	CTG	AGC
57.	FGS57	CGCGCGC	ACATG	TCT	CCC	TGG	GCT	CCC	CTG	AGC	TCC
58.	FGS58	CGCGCGC	ACATG	TCT	TGG	GCT	CCC	CTG	AGC	TCC	TGC
59.	FGS59	CGCGCGC	ACATG	TCT	GCT	CCC	CTG	AGC	TCC	TGC	CCC
60.	FGS60	CGCGCGC	ACATG	TCT	CCC	CTG	AGC	TCC	TGC	CCC	AGC
61.	FGS61	CGCGCGC	ACATG	TCT	CTG	AGC	TCC	TGC	CCC	AGC	CAG
62.	FGS62	CGCGCGC	ACATG	TCT	AGC	TCC	TGC	CCC	AGC	CAG	GCC
63.	FGS63	CGCGCGC	ACATG	TCT	TCC	TGC	CCC	AGC	CAG	GCC	CTG
64.	FGS64	CGCGCGC	ACATG	TCT	TGC	CCC	AGC	CAG	GCC	CTG	CAG
65.	FGS65	CGCGCGC	ACATG	TCT	CCC	AGC	CAG	GCC	CTG	CAG	CTG
66.	FGS66	CGCGCGC	ACATG	TCT	AGC	CAG	GCC	CTG	CAG	CTG	GCA
67.	FGS67	CGCGCGC	ACATG	TCT	CAG	GCC	CTG	CAG	CTG	GCA	GGC
68.	FGS68	CGCGCGC	ACATG	TCT	GCC	CTG	CAG	CTG	GCA	GGC	TGC
69.	FGS69	CGCGCGC	ACATG	TCT	CTG	CAG	CTG	GCA	GGC	TGC	TTG
70.	FGS70	CGCGCGC	ACATG	TCT	CAG	CTG	GCA	GGC	TGC	TTG	AGC
71.	FGS71	CGCGCGC	ACATG	TCT	CTG	GCA	GGC	TGC	TTG	AGC	CAA
72.	FGS72	CGCGCGC	ACATG	TCT	GCA	GGC	TGC	TTG	AGC	CAA	CTC
73.	FGS73	CGCGCGC	ACATG	TCT	GGC	TGC	TTG	AGC	CAA	CTC	CAT
74.	FGS74	CGCGCGC	ACATG	TCT	TGC	TTG	AGC	CAA	CTC	CAT	AGC
75.	FGS75	CGCGCGC	ACATG	TCT	TTG	AGC	CAA	CTC	CAT	AGC	GGC
76.	FGS76	CGCGCGC	ACATG	TCT	AGC	CAA	CTC	CAT	AGC	GGC	CTT
77.	FGS77	CGCGCGC	ACATG	TCT	CAA	CTC	CAT	AGC	GGC	CTT	TTC
78.	FGS78	CGCGCGC	ACATG	TCT	TTC	CAT	AGC	GGC	CTT	TTC	CTC
79.	FGS79	CGCGCGC	ACATG	TCT	CAT	AGC	GGC	CTT	TTC	CTC	TAC

80.	FGS80	CGCGCGC	ACATG	TCT	AGC	GGC	CTT	TTC	CTC	TAC	CAG
81.	FGS81	CGCGCGC	ACATG	TCT	GGC	CTT	TTC	CTC	TAC	CAG	GGG
82.	FGS82	CGCGCGC	ACATG	TCT	CTT	TTC	CTC	TAC	CAG	GGG	CTC
83.	FGS83	CGCGCGC	ACATG	TCT	TTC	CTC	TAC	CAG	GGG	CTC	CTG
84.	FGS84	CGCGCGC	ACATG	TCT	CTC	TAC	CAG	GGG	CTC	CTG	CAG
85.	FGS85	CGCGCGC	ACATG	TCT	TAC	CAG	GGG	CTC	CTG	CAG	GCC
86.	FGS86	CGCGCGC	ACATG	TCT	CAG	GGG	CTC	CTG	CAG	GCC	CTG
87.	FGS87	CGCGCGC	ACATG	TCT	GGG	CTC	CTG	CAG	GCC	CTG	GAA
88.	FGS88	CGCGCGC	ACATG	TCT	CTC	CTG	CAG	GCC	CTG	GAA	GGG
89.	FGS89	CGCGCGC	ACATG	TCT	CTG	CAG	GCC	CTG	GAA	GGG	ATA
90.	FGS90	CGCGCGC	ACATG	TCT	CAG	GCC	CTG	GAA	GGG	ATA	TCC
91.	FGS91	CGCGCGC	ACATG	TCT	GCC	CTG	GAA	GGG	ATA	TCC	CCC
92.	FGS92	CGCGCGC	ACATG	TCT	CTG	GAA	GGG	ATA	TCC	CCC	GAG
93.	FGS93	CGCGCGC	ACATG	TCT	GAA	GGG	ATA	TCC	CCC	GAG	TTG
94.	FGS94	CGCGCGC	ACATG	TCT	GGG	ATA	TCC	CCC	GAG	TTG	GGT
95.	FGS95	CGCGCGC	ACATG	TCT	ATA	TCC	CCC	GAG	TTG	GGT	CCC
96.	FGS96	CGCGCGC	ACATG	TCT	TCC	CCC	GAG	TTG	GGT	CCC	ACC
97.	FGS97	CGCGCGC	ACATG	TCT	CCC	GAG	TTG	GGT	CCC	ACC	TTG
98.	FGS98	CGCGCGC	ACATG	TCT	GAG	TTG	GGT	CCC	ACC	TTG	GAC
99.	FGS99	CGCGCGC	ACATG	TCT	TTG	GGT	CCC	ACC	TTG	GAC	ACA
100.	FGS100	CGCGCGC	ACATG	TCT	GGT	CCC	ACC	TTG	GAC	ACA	CTG
101.	FGS101	CGCGCGC	ACATG	TCT	CCC	ACC	TTG	GAC	ACA	CTG	CAG
102.	FGS102	CGCGCGC	ACATG	TCT	ACC	TTG	GAC	ACA	CTG	CAG	CTG
103.	FGS103	CGCGCGC	ACATG	TCT	TTG	GAC	ACA	CTG	CAG	CTG	GAC
104.	FGS104	CGCGCGC	ACATG	TCT	GAC	ACA	CTG	CAG	CTG	GAC	GTC
105.	FGS105	CGCGCGC	ACATG	TCT	ACA	CTG	CAG	CTG	GAC	GTC	GCC
106.	FGS106	CGCGCGC	ACATG	TCT	CTG	CAG	CTG	GAC	GTC	GCC	GAC
107.	FGS107	CGCGCGC	ACATG	TCT	CAG	CTG	GAC	GTC	GCC	GAC	TTT
108.	FGS108	CGCGCGC	ACATG	TCT	CTG	GAC	GTC	GCC	GAC	TTT	GCC
109.	FGS109	CGCGCGC	ACATG	TCT	GAC	GTC	GCC	GAC	TTT	GCC	ACC
110.	FGS110	CGCGCGC	ACATG	TCT	GTC	GCC	GAC	TTT	GCC	ACC	ACC
111.	FGS111	CGCGCGC	ACATG	TCT	GCC	GAC	TTT	GCC	ACC	ACC	ATC
112.	FGS112	CGCGCGC	ACATG	TCT	GAC	TTT	GCC	ACC	ACC	ATC	TGG
113.	FGS113	CGCGCGC	ACATG	TCT	TTT	GCC	ACC	ACC	ATC	TGG	CAG
114.	FGS114	CGCGCGC	ACATG	TCT	GCC	ACC	ACC	ATC	TGG	CAG	CAG
115.	FGS115	CGCGCGC	ACATG	TCT	ACC	ACC	ATC	TGG	CAG	CAG	ATG
116.	FGS116	CGCGCGC	ACATG	TCT	ACC	ATC	TGG	CAG	CAG	ATG	GAA
117.	FGS117	CGCGCGC	ACATG	TCT	ATC	TGG	CAG	CAG	ATG	GAA	GAA
118.	FGS118	CGCGCGC	ACATG	TCT	TGG	CAG	CAG	ATG	GAA	GAA	CTG
119.	FGS119	CGCGCGC	ACATG	TCT	CAG	CAG	ATG	GAA	GAA	CTG	GGA
120.	FGS120	CGCGCGC	ACATG	TCT	CAG	ATG	GAA	GAA	CTG	GGA	ATG
121.	FGS121	CGCGCGC	ACATG	TCT	ATG	GAA	GAA	CTG	GGA	ATG	GCC
122.	FGS122	CGCGCGC	ACATG	TCT	GAA	GAA	CTG	GGA	ATG	GCC	CCT
123.	FGS123	CGCGCGC	ACATG	TCT	GAA	CTG	GGA	ATG	GCC	CCT	GCC
124.	FGS124	CGCGCGC	ACATG	TCT	CTG	GGA	ATG	GCC	CCT	GCC	CTG
125.	FGS125	CGCGCGC	ACATG	TCT	GGA	ATG	GCC	CCT	GCC	CTG	CAG
126.	FGS126	CGCGCGC	ACATG	TCT	ATG	GCC	CCT	GCC	CTG	CAG	CCC
127.	FGS127	CGCGCGC	ACATG	TCT	GCC	CCT	GCC	CTG	CAG	CCC	ACC
128.	FGS128	CGCGCGC	ACATG	TCT	CCT	GCC	CTG	CAG	CCC	ACC	CAG
129.	FGS129	CGCGCGC	ACATG	TCT	GCC	CTG	CAG	CCC	ACC	CAG	GGT
130.	FGS130	CGCGCGC	ACATG	TCT	CTG	CAG	CCC	ACC	CAG	GGT	GCC
131.	FGS131	CGCGCGC	ACATG	TCT	CAG	CCC	ACC	CAG	GGT	GCC	ATG
132.	FGS132	CGCGCGC	ACATG	TCT	CCC	ACC	CAG	GGT	GCC	ATG	CCG
133.	FGS133	CGCGCGC	ACATG	TCT	ACC	CAG	GGT	GCC	ATG	CCG	GCC
134.	FGS134	CGCGCGC	ACATG	TCT	CAG	GGT	GCC	ATG	CCG	GCC	TTC
135.	FGS135	CGCGCGC	ACATG	TCT	GGT	GCC	ATG	CCG	GCC	TTC	GCC
136.	FGS136	CGCGCGC	ACATG	TCT	GCC	ATG	CCG	GCC	TTC	GCC	TCT
137.	FGS137	CGCGCGC	ACATG	TCT	ATG	CCG	GCC	TTC	GCC	TCT	GCT
138.	FGS138	CGCGCGC	ACATG	TCT	CCG	GCC	TTC	GCC	TCT	GCT	TTC
139.	FGS139	CGCGCGC	ACATG	TCT	GCC	TTC	GCC	TCT	GCT	TTC	CAG
140.	FGS140	CGCGCGC	ACATG	TCT	TTC	GCC	TCT	GCT	TTC	CAG	CGC
141.	FGS141	CGCGCGC	ACATG	TCT	GCC	TCT	GCT	TTC	CAG	CGC	CGG
142.	FGS142	CGCGCGC	ACATG	TCT	TCT	GCT	TTC	CAG	CGC	CGG	GCA
143.	FGS143	CGCGCGC	ACATG	TCT	GCT	TTC	CAG	CGC	CGG	GCA	GGA
144.	FGS144	CGCGCGC	ACATG	TCT	TTC	CAG	CGC	CGG	GCA	GGA	GGG
145.	FGS145	CGCGCGC	ACATG	TCT	CAG	CGC	CGG	GCA	GGA	GGG	GTC
146.	FGS146	CGCGCGC	ACATG	TCT	CGC	CGG	GCA	GGA	GGG	GTC	CTG
147.	FGS147	CGCGCGC	ACATG	TCT	CGG	GCA	GGA	GGG	GTC	CTG	GTT
148.	FGS148	CGCGCGC	ACATG	TCT	GCA	GGA	GGG	GTC	CTG	GTT	GCT
149.	FGS149	CGCGCGC	ACATG	TCT	GGA	GGG	GTC	CTG	GTT	GCT	AGC
150.	FGS150	CGCGCGC	ACATG	TCT	GGG	GTC	CTG	GTT	GCT	AGC	CAT
151.	FGS151	CGCGCGC	ACATG	TCT	GTC	CTG	GTT	GCT	AGC	CAT	CTG
152.	FGS152	CGCGCGC	ACATG	TCT	CTG	GTT	GCT	AGC	CAT	CTG	CAG
153.	FGS153	CGCGCGC	ACATG	TCT	GTT	GCT	AGC	CAT	CTG	CAG	AGC
154	FGS154	CGCGCGC	ACATG	TCT	GCT	AGC	CAT	CTG	CAG	AGC	TTC
155.	FGS155	CGCGCGC	ACATG	TCT	AGC	CAT	CTG	CAG	AGC	TTC	CTG
156	FGS156	CGCGCGC	ACATG	TCT	CAT	CTG	CAG	AGC	TTC	CTG	GAG
157.	FGS157	CGCGCGC	ACATG	TCT	CTG	CAG	AGC	TTC	CTG	GAG	GTG
158.	FGS158	CGCGCGC	ACATG	TCT	CAG	AGC	TTC	CTG	GAG	GTG	TCG
159.	FGS159	CGCGCGC	ACATG	TCT	AGC	TTC	CTG	GAG	GTG	TCG	TAC
160.	FGS160	CGCGCGC	ACATG	TCT	TTC	CTG	GAG	GTG	TCG	TAC	CGC
161.	FGS161	CGCGCGC	ACATG	TCT	CTG	GAG	GTG	TCG	TAC	CGC	GTT
162.	FGS162	CGCGCGC	ACATG	TCT	GAG	GTG	TCG	TAC	CGC	GTT	CTA
163.	FGS163	CGCGCGC	ACATG	TCT	GTG	TCG	TAC	CGC	GTT	CTA	CGC
164.	FGS164	CGCGCGC	ACATG	TCT	TCG	TAC	CGC	GTT	CTA	CGC	CAC
165.	FGS165	CGCGCGC	ACATG	TCT	TAC	CGC	GTT	CTA	CGC	CAC	GTT

166.	FGS166	CGCGCGC	ACATG	TCT	CGC	GTT	CTA	CGC	CAC	CTT	GCG
167.	FGS167	CGCGCGC	ACATG	TCT	GTT	CTA	CGC	CAC	CTT	GCG	CAG
168.	FGS168	CGCGCGC	ACATG	TCT	CTA	CGC	CAC	CTT	GCG	CAG	CCC
169.	FGS169A	CGCGCGC	ACATG	TCT	CGC	CAC	CTT	GCG	CAG	CCC	GA'C
170.	FGS170A	CGCGCGC	ACATG	TCT	CAC	CTT	GCG	CAG	CCC	GA'C	ATG
171.	FGS171A	CGCGCGC	ACATG	TCT	CTT	GCG	CAG	CCC	GA'C	ATG	GCT
172.	FGS172A	CGCGCGC	ACATG	TCT	GCG	CAG	CCC	GA'C	ATG	GCT	ACA
173.	FGS173A	CGCGCGC	ACATG	TCT	CAG	CCC	GA'C	ATG	GCT	ACA	CCA
174.	FGS174A	CGCGCGC	ACATG	TCT	CCC	GA'C	ATG	GCT	ACA	CCA	TTG
175.	FGS169B	CGCGCGC	ACATG	TCT	CGC	CAC	CTT	GCG	CAG	CCC	A'CT
176.	FGS170B	CGCGCGC	ACATG	TCT	CAC	CTT	GCG	CAG	CCC	A'CT	AGT
177.	FGS171B	CGCGCGC	ACATG	TCT	CTT	GCG	CAG	CCC	A'CT	AGT	CAT
178.	FGS172B	CGCGCGC	ACATG	TCT	GCG	CAG	CCC	A'CT	AGT	CAT	CCA
179.	FGS173B	CGCGCGC	ACATG	TCT	CAG	CCC	A'CT	AGT	CAT	CCA	CCT
180.	FGS174B	CGCGCGC	ACATG	TCT	CCC	A'CT	AGT	CAT	CCA	CCT	ATG
181.	FGS169C	CGCGCGC	ACATG	TCT	CGC	CAC	CTT	GCG	CAG	CCC	GGC
182.	FGS170C	CGCGCGC	ACATG	TCT	CAC	CTT	GCG	CAG	CCC	GGC	GGC
183.	FGS171C	CGCGCGC	ACATG	TCT	CTT	GCG	CAG	CCC	GGC	GGC	GGC
184.	FGS172C	CGCGCGC	ACATG	TCT	GCG	CAG	CCC	GGC	GGC	GGC	TCT
185.	FGS173C	CGCGCGC	ACATG	TCT	CAG	CCC	GGC	GGC	GGC	TCT	GA'C
186.	FGS174C	CGCGCGC	ACATG	TCT	CCC	GGC	GGC	GGC	TCT	GA'C	ATG
187.	RGS0A	TATATAT	GCGGCCGC	AGC	CAT	GTC	GGG	CTG	CGC	AAG	
188.	RGS0B	TATATAT	GCGGCCGC	AGC	CAT	GTC	ACG	CGT	ACG	ATT	
189.	RGS0C	TATATAT	GCGGCCGC	AGC	CAT	GTC	AGA	GCC	GCC	GCC	
190.	RGS1A	TATATAT	GCGGCCGC	TGT	AGC	CAT	GTC	GGG	CTG	CGC	
191.	RGS1B	TATATAT	GCGGCCGC	TGT	AGC	CAT	GTC	ACG	CGT	ACG	
192.	RGS1C	TATATAT	GCGGCCGC	TGT	AGC	CAT	GTC	AGA	GCC	GCC	
193.	RGS2A	TATATAT	GCGGCCGC	TGG	TGT	AGC	CAT	GTC	GGG	CTG	
194.	RGS2B	TATATAT	GCGGCCGC	TGG	TGT	AGC	CAT	GTC	ACG	CGT	
195.	RGS2C	TATATAT	GCGGCCGC	TGG	TGT	AGC	CAT	GTC	AGA	GCC	
196.	RGS3A	TATATAT	GCGGCCGC	CAA	TGG	TGT	AGC	CAT	GTC	GGG	
197.	RGS3B	TATATAT	GCGGCCGC	CAA	TGG	TGT	AGC	CAT	GTC	ACG	
198.	RGS3C	TATATAT	GCGGCCGC	CAA	TGG	TGT	AGC	CAT	GTC	AGA	
199.	RGS4	TATATAT	GCGGCCGC	GCC	CAA	TGG	TGT	AGC	CAT	GTC	
200.	RGS5	TATATAT	GCGGCCGC	AGG	GCC	CAA	TGG	TGT	AGC	CAT	
201.	RGS6	TATATAT	GCGGCCGC	GGC	AGG	GCC	CAA	TGG	TGT	AGC	
202.	RGS7	TATATAT	GCGGCCGC	GCT	GGC	AGG	GCC	CAA	TGG	TGT	
203.	RGS8	TATATAT	GCGGCCGC	GGA	GCT	GGC	AGG	GCC	CAA	TGG	
204.	RGS9	TATATAT	GCGGCCGC	CAG	GGA	GCT	GGC	AGG	GCC	CAA	
205.	RGS10	TATATAT	GCGGCCGC	GGG	CAG	GGA	GCT	GGC	AGG	GCC	
206.	RGS11	TATATAT	GCGGCCGC	CTG	GGG	CAG	GGA	GCT	GGC	AGG	
207.	RGS12	TATATAT	GCGGCCGC	GCT	CTG	GGG	CAG	GGA	GCT	GGC	
208.	RGS13	TATATAT	GCGGCCGC	GAA	GCT	CTG	GGG	CAG	GGA	GCT	
209.	RGS14	TATATAT	GCGGCCGC	CAG	GAA	GCT	CTG	GGG	CAG	GGA	
210.	RGS15	TATATAT	GCGGCCGC	GAG	CAG	GAA	GCT	CTG	GGG	CAG	
211.	RGS16	TATATAT	GCGGCCGC	CTT	GAG	CAG	GAA	GCT	CTG	GGG	
212.	RGS17	TATATAT	GCGGCCGC	AGA	CTT	GAG	CAG	GAA	GCT	CTG	
213.	RGS18	TATATAT	GCGGCCGC	TAA	AGA	CTT	GAG	CAG	GAA	GCT	
214.	RGS19	TATATAT	GCGGCCGC	CTC	TAA	AGA	CTT	GAG	CAG	GAA	
215.	RGS20	TATATAT	GCGGCCGC	TTG	CTC	TAA	AGA	CTT	GAG	CAG	
216.	RGS21	TATATAT	GCGGCCGC	CAC	TTG	CTC	TAA	AGA	CTT	GAG	
217.	RGS22	TATATAT	GCGGCCGC	CCT	CAC	TTG	CTC	TAA	AGA	CTT	
218.	RGS23	TATATAT	GCGGCCGC	CTT	CCT	CAC	TTG	CTC	TAA	AGA	
219.	RGS24	TATATAT	GCGGCCGC	GAT	CTT	CCT	CAC	TTG	CTC	TAA	
220.	RGS25	TATATAT	GCGGCCGC	CTG	GAT	CTT	CCT	CAC	TTG	CTC	
221.	RGS26	TATATAT	GCGGCCGC	GCC	CTG	GAT	CTT	CCT	CAC	TTG	
222.	RGS27	TATATAT	GCGGCCGC	ATC	GCC	CTG	GAT	CTT	CCT	CAC	
223.	RGS28	TATATAT	GCGGCCGC	GCC	ATC	GCC	CTG	GAT	CTT	CCT	
224.	RGS29	TATATAT	GCGGCCGC	TGC	GCC	ATC	GCC	CTG	GAT	CTT	
225.	RGS30	TATATAT	GCGGCCGC	CGC	TGC	GCC	ATC	GCC	CTG	GAT	
226.	RGS31	TATATAT	GCGGCCGC	GAG	CGC	TGC	GCC	ATC	GCC	CTG	
227.	RGS32	TATATAT	GCGGCCGC	CTG	GAG	CGC	TGC	GCC	ATC	GCC	
228.	RGS33	TATATAT	GCGGCCGC	CTC	CTG	GAG	CGC	TGC	GCC	ATC	
229.	RGS34	TATATAT	GCGGCCGC	CTT	CTC	CTG	GAG	CGC	TGC	GCC	
230.	RGS35	TATATAT	GCGGCCGC	CAG	CTT	CTC	CTG	GAG	CGC	TGC	
231.	RGS36	TATATAT	GCGGCCGC	ACA	CAG	CTT	CTC	CTG	GAG	CGC	
232.	RGS37	TATATAT	GCGGCCGC	GGC	ACA	CAG	CTT	CTC	CTG	GAG	
233.	RGS38	TATATAT	GCGGCCGC	GGT	GGC	ACA	CAG	CTT	CTC	CTG	
234.	RGS39	TATATAT	GCGGCCGC	GTA	GGT	GGC	ACA	CAG	CTT	CTC	
235.	RGS40	TATATAT	GCGGCCGC	CTT	GTA	GGT	GGC	ACA	CAG	CTT	
236.	RGS41	TATATAT	GCGGCCGC	CAG	CTT	GTA	GGT	GGC	ACA	CAG	
237.	RGS42	TATATAT	GCGGCCGC	GCA	CAG	CTT	GTA	GGT	GGC	ACA	
238.	RGS43	TATATAT	GCGGCCGC	GTG	GCA	CAG	CTT	GTA	GGT	GGC	
239.	RGS44	TATATAT	GCGGCCGC	GGG	GTG	GCA	CAG	CTT	GTA	GGT	
240.	RGS45	TATATAT	GCGGCCGC	CTC	GGG	GTG	GCA	CAG	CTT	GTA	
241.	RGS46	TATATAT	GCGGCCGC	CTC	CTC	GGG	GTG	GCA	CAG	CTT	
242.	RGS47	TATATAT	GCGGCCGC	CAG	CTC	CTC	GGG	GTG	GCA	CAG	
243.	RGS48	TATATAT	GCGGCCGC	CAC	CAG	CTC	CTC	GGG	GTG	GCA	
244.	RGS49	TATATAT	GCGGCCGC	CAG	CAC	CAG	CTC	CTC	GGG	GTG	
245.	RGS50	TATATAT	GCGGCCGC	GAG	CAG	CAC	CAG	CTC	CTC	GGG	
246.	RGS51	TATATAT	GCGGCCGC	TCC	GAG	CAG	CAC	CAG	CTC	CTC	
247.	RGS52	TATATAT	GCGGCCGC	GTG	TCC	GAG	CAG	CAC	CAG	CTC	
248.	RGS53	TATATAT	GCGGCCGC	AGA	GTG	TCC	GAG	CAG	CAC	CAG	
249.	RGS54	TATATAT	GCGGCCGC	CAG	AGA	GTG	TCC	GAG	CAG	CAC	
250.	RGS55	TATATAT	GCGGCCGC	GCC	CAG	AGA	GTG	TCC	GAG	CAG	
251.	RGS56	TATATAT	GCGGCCGC	GAT	GCC	CAG	AGA	GTG	TCC	GAG	

252.	RGS57	TATATAT	GCGGCCGC	GGG	GAT	GCC	CAG	AGA	GTG	TCC
253.	RGS58	TATATAT	GCGGCCGC	CCA	GGG	GAT	GCC	CAG	AGA	GTG
254.	RGS59	TATATAT	GCGGCCGC	AGC	CCA	GGG	GAT	GCC	CAG	AGA
255.	RGS60	TATATAT	GCGGCCGC	GGG	AGC	CCA	GGG	GAT	GCC	CAG
256.	RGS61	TATATAT	GCGGCCGC	CAG	GGG	AGC	CCA	GGG	GAT	GCC
257.	RGS62	TATATAT	GCGGCCGC	GCT	CAG	GGG	AGC	CCA	GGG	GAT
258.	RGS63	TATATAT	GCGGCCGC	GGA	GCT	CAG	GGG	AGC	CCA	GGG
259.	RGS64	TATATAT	GCGGCCGC	GCA	GGA	GCT	CAG	GGG	AGC	CCA
260.	RGS65	TATATAT	GCGGCCGC	GGG	GCA	GGA	GCT	CAG	GGG	AGC
261.	RGS66	TATATAT	GCGGCCGC	GCT	GGG	GCA	GGA	GCT	CAG	GGG
262.	RGS67	TATATAT	GCGGCCGC	CTG	GCT	GGG	GCA	GGA	GCT	CAG
263.	RGS68	TATATAT	GCGGCCGC	GGC	CTG	GCT	GGG	GCA	GGA	GCT
264.	RGS69	TATATAT	GCGGCCGC	CAG	GGC	CTG	GCT	GGG	GCA	GGA
265.	RGS70	TATATAT	GCGGCCGC	CTG	CAG	GGC	CTG	GCT	GGG	GCA
266.	RGS71	TATATAT	GCGGCCGC	CAG	CTG	CAG	GGC	CTG	GCT	GGG
267.	RGS72	TATATAT	GCGGCCGC	TGC	CAG	CTG	CAG	GGC	CTG	GCT
268.	RGS73	TATATAT	GCGGCCGC	GCC	TGC	CAG	CTG	CAG	GGC	CTG
269.	RGS74	TATATAT	GCGGCCGC	GCA	GCC	TGC	CAG	CTG	CAG	GGC
270.	RGS75	TATATAT	GCGGCCGC	CAA	GCA	GCC	TGC	CAG	CTG	CAG
271.	RGS76	TATATAT	GCGGCCGC	GCT	CAA	GCA	GCC	TGC	CAG	CTG
272.	RGS77	TATATAT	GCGGCCGC	GTT	GCT	CAA	GCA	GCC	TGC	CAG
273.	RGS78	TATATAT	GCGGCCGC	GAG	GTT	GCT	CAA	GCA	GCC	TGC
274.	RGS79	TATATAT	GCGGCCGC	ATG	GAG	GTT	GCT	CAA	GCA	GCC
275.	RGS80	TATATAT	GCGGCCGC	GCT	ATG	GAG	GTT	GCT	CAA	GCA
276.	RGS81	TATATAT	GCGGCCGC	GCC	GCT	ATG	GAG	GTT	GCT	CAA
277.	RGS82	TATATAT	GCGGCCGC	AAG	GCC	GCT	ATG	GAG	GTT	GCT
278.	RGS83	TATATAT	GCGGCCGC	GAA	AAG	GCC	GCT	ATG	GAG	GTT
279.	RGS84	TATATAT	GCGGCCGC	GAG	GAA	AAG	GCC	GCT	ATG	GAG
280.	RGS85	TATATAT	GCGGCCGC	GTA	GAG	GAA	AAG	GCC	GCT	ATG
281.	RGS86	TATATAT	GCGGCCGC	CTG	GTA	GAG	GAA	AAG	GCC	GCT
282.	RGS87	TATATAT	GCGGCCGC	CCC	CTG	GTA	GAG	GAA	AAG	GCC
283.	RGS88	TATATAT	GCGGCCGC	GAG	CCC	CTG	GTA	GAG	GAA	AAG
284.	RGS89	TATATAT	GCGGCCGC	CAG	GAG	CCC	CTG	GTA	GAG	GAA
285.	RGS90	TATATAT	GCGGCCGC	CTG	CAG	GAG	CCC	CTG	GTA	GAG
286.	RGS91	TATATAT	GCGGCCGC	GGC	CTG	CAG	GAG	CCC	CTG	GTA
287.	RGS92	TATATAT	GCGGCCGC	CAG	GGC	CTG	CAG	GAG	CCC	CTG
288.	RGS93	TATATAT	GCGGCCGC	TTC	CAG	GGC	CTG	CAG	GAG	CCC
289.	RGS94	TATATAT	GCGGCCGC	CCC	TTC	CAG	GGC	CTG	CAG	GAG
290.	RGS95	TATATAT	GCGGCCGC	TAT	CCC	TTC	CAG	GGC	CTG	CAG
291.	RGS96	TATATAT	GCGGCCGC	GGA	TAT	CCC	TTC	CAG	GGC	CTG
292.	RGS97	TATATAT	GCGGCCGC	GGG	GGA	TAT	CCC	TTC	CAG	GGC
293.	RGS98	TATATAT	GCGGCCGC	CTC	GGG	GGA	TAT	CCC	TTC	CAG
294.	RGS99	TATATAT	GCGGCCGC	CAA	CTC	GGG	GGA	TAT	CCC	TTC
295.	RGS100	TATATAT	GCGGCCGC	ACC	CAA	CTC	GGG	GGA	TAT	CCC
296.	RGS101	TATATAT	GCGGCCGC	GGG	ACC	CAA	CTC	GGG	GGA	TAT
297.	RGS102	TATATAT	GCGGCCGC	GGT	GGG	ACC	CAA	CTC	GGG	GGA
298.	RGS103	TATATAT	GCGGCCGC	CAA	GGT	GGG	ACC	CAA	CTC	GGG
299.	RGS104	TATATAT	GCGGCCGC	GTC	CAA	GGT	GGG	ACC	CAA	CTC
300.	RGS105	TATATAT	GCGGCCGC	TGT	GTC	CAA	GGT	GGG	ACC	CAA
301.	RGS106	TATATAT	GCGGCCGC	CAG	TGT	GTC	CAA	GGT	GGG	ACC
302.	RGS107	TATATAT	GCGGCCGC	CTG	CAG	TGT	GTC	CAA	GGT	GGG
303.	RGS108	TATATAT	GCGGCCGC	CAG	CTG	CAG	TGT	GTC	CAA	GGT
304.	RGS109	TATATAT	GCGGCCGC	GTC	CAG	CTG	CAG	TGT	GTC	CAA
305.	RGS110	TATATAT	GCGGCCGC	GAC	GTC	CAG	CTG	CAG	TGT	GTC
306.	RGS111	TATATAT	GCGGCCGC	GGC	GAC	GTC	CAG	CTG	CAG	TGT
307.	RGS112	TATATAT	GCGGCCGC	GTC	GGC	GAC	GTC	CAG	CTG	CAG
308.	RGS113	TATATAT	GCGGCCGC	AAA	GTC	GGC	GAC	GTC	CAG	CTG
309.	RGS114	TATATAT	GCGGCCGC	GGC	AAA	GTC	GGC	GAC	GTC	CAG
310.	RGS115	TATATAT	GCGGCCGC	GGT	GGC	AAA	GTC	GGC	GAC	GTC
311.	RGS116	TATATAT	GCGGCCGC	GGT	GGT	GGC	AAA	GTC	GGC	GAC
312.	RGS117	TATATAT	GCGGCCGC	GAT	GGT	GGT	GGC	AAA	GTC	GGC
313.	RGS118	TATATAT	GCGGCCGC	CCA	GAT	GGT	GGT	GGC	AAA	GTC
314.	RGS119	TATATAT	GCGGCCGC	CTG	CCA	GAT	GGT	GGT	GGC	AAA
315.	RGS120	TATATAT	GCGGCCGC	CTG	CTG	CCA	GAT	GGT	GGT	GGC
316.	RGS121	TATATAT	GCGGCCGC	CAT	CTG	CTG	CCA	GAT	GGT	GGT
317.	RGS122	TATATAT	GCGGCCGC	TTC	CAT	CTG	CTG	CCA	GAT	GGT
318.	RGS123	TATATAT	GCGGCCGC	TTC	TTC	CAT	CTG	CTG	CCA	GAT
319.	RGS124	TATATAT	GCGGCCGC	CAG	TTC	TTC	CAT	CTG	CTG	CCA
320.	RGS125	TATATAT	GCGGCCGC	TCC	CAG	TTC	TTC	CAT	CTG	CTG
321.	RGS126	TATATAT	GCGGCCGC	CAT	TCC	CAG	TTC	TTC	CAT	CTG
322.	RGS127	TATATAT	GCGGCCGC	GGC	CAT	TCC	CAG	TTC	TTC	CAT
323.	RGS128	TATATAT	GCGGCCGC	AGG	GGC	CAT	TCC	CAG	TTC	TTC
324.	RGS129	TATATAT	GCGGCCGC	GGC	AGG	GGC	CAT	TCC	CAG	TTC
325.	RGS130	TATATAT	GCGGCCGC	CAG	GGC	AGG	GGC	CAT	TCC	CAG
326.	RGS131	TATATAT	GCGGCCGC	CTG	CAG	GGC	AGG	GGC	CAT	TCC
327.	RGS132	TATATAT	GCGGCCGC	GGG	CTG	CAG	GGC	AGG	GGC	CAT
328.	RGS133	TATATAT	GCGGCCGC	GGT	GGG	CTG	CAG	GGC	AGG	GGC
329.	RGS134	TATATAT	GCGGCCGC	CTG	GGT	GGG	CTG	CAG	GGC	AGG
330.	RGS135	TATATAT	GCGGCCGC	ACC	CTG	GGT	GGG	CTG	CAG	GGC
331.	RGS136	TATATAT	GCGGCCGC	GGC	ACC	CTG	GGT	GGG	CTG	CAG
332.	RGS137	TATATAT	GCGGCCGC	CAT	GGC	ACC	CTG	GGT	GGG	CTG
333.	RGS138	TATATAT	GCGGCCGC	CGG	CAT	GGC	ACC	CTG	GGT	GGG
334.	RGS139	TATATAT	GCGGCCGC	GGC	CGG	CAT	GGC	ACC	CTG	GGT
335.	RGS140	TATATAT	GCGGCCGC	GAA	GGC	CGG	CAT	GGC	ACC	CTG
336.	RGS141	TATATAT	GCGGCCGC	GGC	GAA	GGC	CGG	CAT	GGC	ACC
337.	RGS142	TATATAT	GCGGCCGC	AGA	GGC	GAA	GGC	CGG	CAT	GGC

338.	RGS143	TATATAT	GCGGCCGC	AGC	AGA	GGC	GAA	GGC	CGG	CAT
339.	RGS144	TATATAT	GCGGCCGC	GAA	AGC	AGA	GGC	GAA	GGC	CGG
340.	RGS145	TATATAT	GCGGCCGC	CTG	GAA	AGC	AGA	GGC	GAA	GGC
341.	RGS146	TATATAT	GCGGCCGC	GCG	CTG	GAA	AGC	AGA	GGC	GAA
342.	RGS147	TATATAT	GCGGCCGC	CCG	GCG	CTG	GAA	AGC	AGA	GGC
343.	RGS148	TATATAT	GCGGCCGC	TGC	CCG	GCG	CTG	GAA	AGC	AGA
344.	RGS149	TATATAT	GCGGCCGC	TCC	TGC	CCG	GCG	CTG	GAA	AGC
345.	RGS150	TATATAT	GCGGCCGC	CCC	TCC	TGC	CCG	GCG	CTG	GAA
346.	RGS151	TATATAT	GCGGCCGC	GAC	CCC	TCC	TGC	CCG	GCG	CTG
347.	RGS152	TATATAT	GCGGCCGC	CAG	GAC	CCC	TCC	TGC	CCG	GCG
348.	RGS153	TATATAT	GCGGCCGC	AAC	CAG	GAC	CCC	TCC	TGC	CCG
349.	RGS154	TATATAT	GCGGCCGC	AGC	AAC	CAG	GAC	CCC	TCC	TGC
350.	RGS155	TATATAT	GCGGCCGC	GCT	AGC	AAC	CAG	GAC	CCC	TCC
351.	RGS156	TATATAT	GCGGCCGC	ATG	GCT	AGC	AAC	CAG	GAC	CCC
352.	RGS157	TATATAT	GCGGCCGC	CAG	ATG	GCT	AGC	AAC	CAG	GAC
353.	RGS158	TATATAT	GCGGCCGC	CTG	CAG	ATG	GCT	AGC	AAC	CAG
354.	RGS159	TATATAT	GCGGCCGC	GCT	CTG	CAG	ATG	GCT	AGC	AAC
355.	RGS160	TATATAT	GCGGCCGC	GAA	GCT	CTG	CAG	ATG	GCT	AGC
356.	RGS161	TATATAT	GCGGCCGC	CAG	GAA	GCT	CTG	CAG	ATG	GCT
357.	RGS162	TATATAT	GCGGCCGC	CTC	CAG	GAA	GCT	CTG	CAG	ATG
358.	RGS163	TATATAT	GCGGCCGC	CAC	CTC	CAG	GAA	GCT	CTG	CAG
359.	RGS164	TATATAT	GCGGCCGC	CGA	CAC	CTC	CAG	GAA	GCT	CTG
360.	RGS165	TATATAT	GCGGCCGC	GTA	CGA	CAC	CTC	CAG	GAA	GCT
361.	RGS166	TATATAT	GCGGCCGC	GCG	GTA	CGA	CAC	CTC	CAG	GAA
362.	RGS167	TATATAT	GCGGCCGC	AAC	GCG	GTA	CGA	CAC	CTC	CAG
363.	RGS168	TATATAT	GCGGCCGC	TAG	AAC	GCG	GTA	CGA	CAC	CTC
364.	RGS169	TATATAT	GCGGCCGC	GCG	TAG	AAC	GCG	GTA	CGA	CAC
365.	RGS170	TATATAT	GCGGCCGC	GTG	GCG	TAG	AAC	GCG	GTA	CGA
366.	RGS171	TATATAT	GCGGCCGC	AAG	GTG	GCG	TAG	AAC	GCG	GTA
367.	RGS172	TATATAT	GCGGCCGC	CGC	AAG	GTG	GCG	TAG	AAC	GCG
368.	RGS173	TATATAT	GCGGCCGC	CTG	CGC	AAG	GTG	GCG	TAG	AAC	GCG

CLAIMS

What is claimed is:

1. A method for making a biologically-active circularly-permuted protein of the formula C^1-L^1-N^1, derived from a parent protein of the formula N^1-C^1, wherein

C^1 is comprised of a segment derived from the carboxy portion of said parent protein;

N^1 is comprised of a segment derived from the amino terminal portion of said parent protein; and

L^1 is a chemical bond or a linker, linking C^1 to the amino terminus of L^1 and carboxy terminus of L^1 to the amino terminus of N^1;

comprising the steps of:

(a) making a series of circularly-permuted genes;

(b) inserting said circularly-permuted genes into a display vector;

(c) expressing said circularly-permuted genes such that the proteins encoded by said genes are presented on the surface of the display vector;

(d) generating a library of display vectors presenting the expressed circularly permuted protein;

(e) affinity-selecting the presenting display vectors with a target protein that can bind a biologically-active circularly-permuted protein;

(f) isolating and analyzing clones of selected display vectors to identify the presented circularly-permuted protein.

2. The method of claim 1 wherein the method of making a series of circularly-permuted genes is selected from the group consisting of making a tandemly-repeated intermediate, total synthesis of a synthetic gene, assembly of a gene from synthetic oligonucleotides, DNA amplification, and limited digestion of a circular

intermediate.

3. The method of claim 1 wherein said display vector is selected from the group consisting of bacteriophage display vectors, bacteria, and baculovirus vectors.

4. The method of claim 3, wherein said presentation vector is a bacteriophage.

5. The method of claim 4, wherein said presentation vector is bacteriophage M13.

6. The method of claim 5, wherein said presentation vector is a bacteriophage M13 gene III vector.

7. The method of claim 1 wherein said method of making a series of circularly permuted genes is a method of making a tandem repeat intermediate.

8 The method of claim 7 wherein said circularly-permuted genes are amplified from the repeat by gene amplification.

9. The method of claim 1 wherein said method of affinity selection comprises the steps consisting of:

(a) binding said presentation display vectors to a target protein;

(b) eluting said display vectors;

(c) amplifying said display vectors; and

(d) biopanning a pool of said amplified display vectors.

10. The method of claim 1 wherein L^1 is a linear peptide linker.

11. The method of claim 1 wherein said the DNA sequence encoding said linker L^1 is selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 368.

12. The method of claim 1 wherein the length of the C^1 in said permutein is shorter than the length of C^1 in said parent protein.

13. The method of claim 1 wherein the length of the N^1 in said permutein is shorter than the length of N^1 in said parent protein.

14. A circularly-permuted protein prepared by the method of claim 1.

15. A circularly-permuted protein of claim 14 comprising the G-CSF receptor agonist domain of a species of mylepoietin (MPO).

Figure 1A

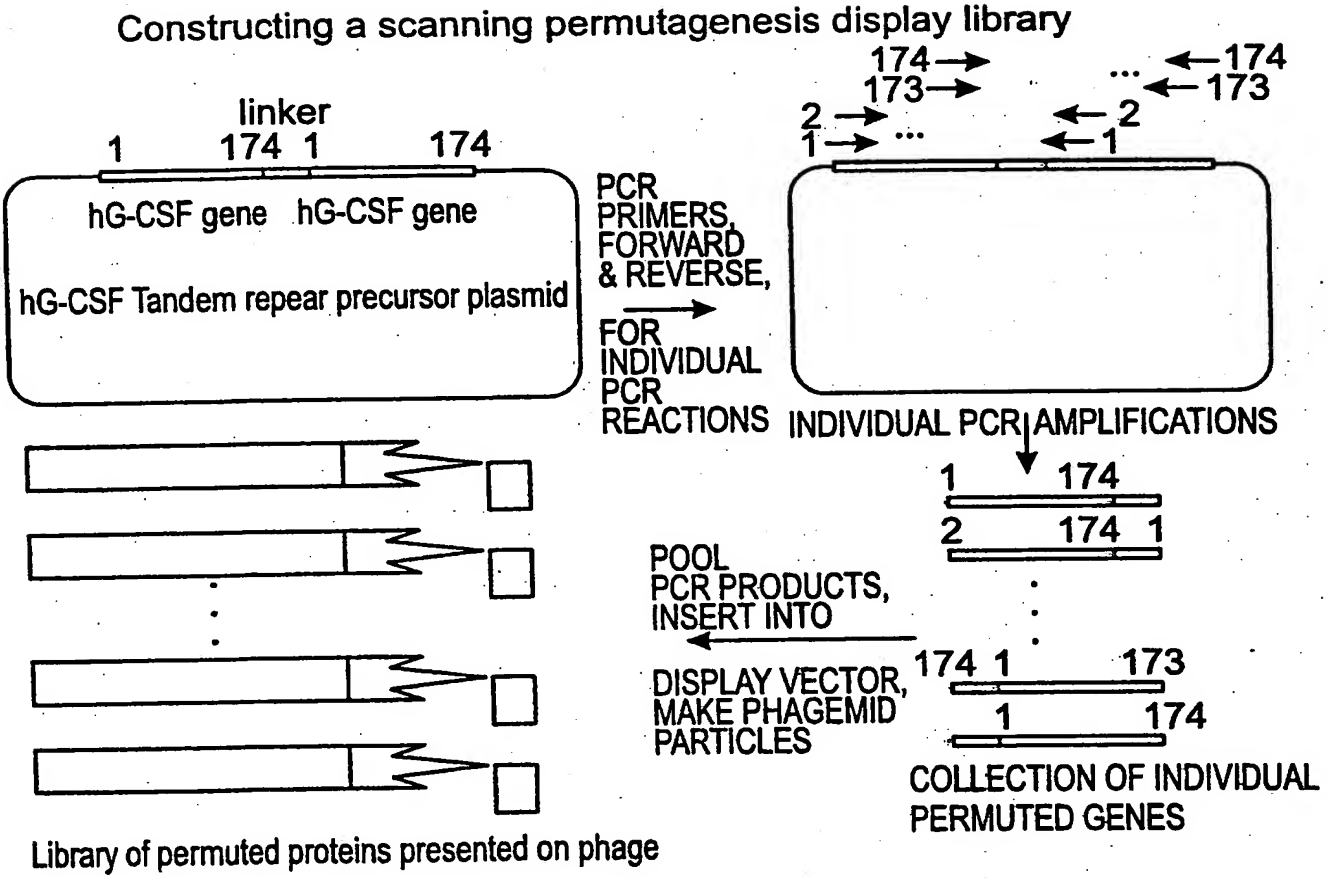

Constructing a scanning permutagenesis display library
linker
1 174 1 174
hG-CSF gene hG-CSF gene
hG-CSF Tandem repear precursor plasmid
PCR PRIMERS, FORWARD & REVERSE,
FOR INDIVIDUAL PCR REACTIONS
174 173 2 1
174 173 2 1
INDIVIDUAL PCR AMPLIFICATIONS
1 174
2 174 1
POOL PCR PRODUCTS, INSERT INTO
DISPLAY VECTOR, MAKE PHAGEMID PARTICLES
174 1 173
1 174
COLLECTION OF INDIVIDUAL PERMUTED GENES
Library of permuted proteins presented on phage

Figure 1B

Screening a display library

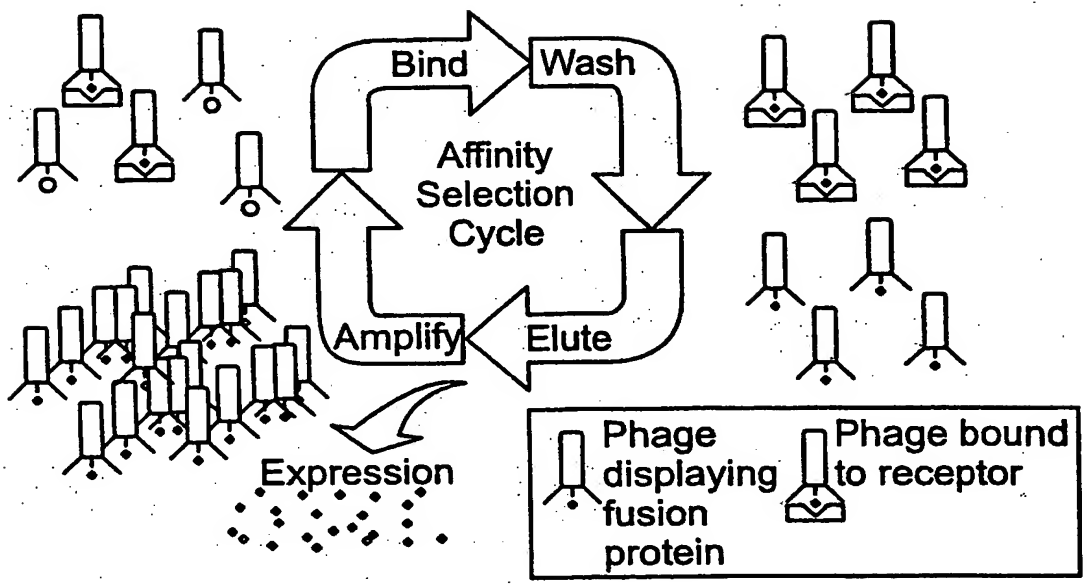


Figure 2

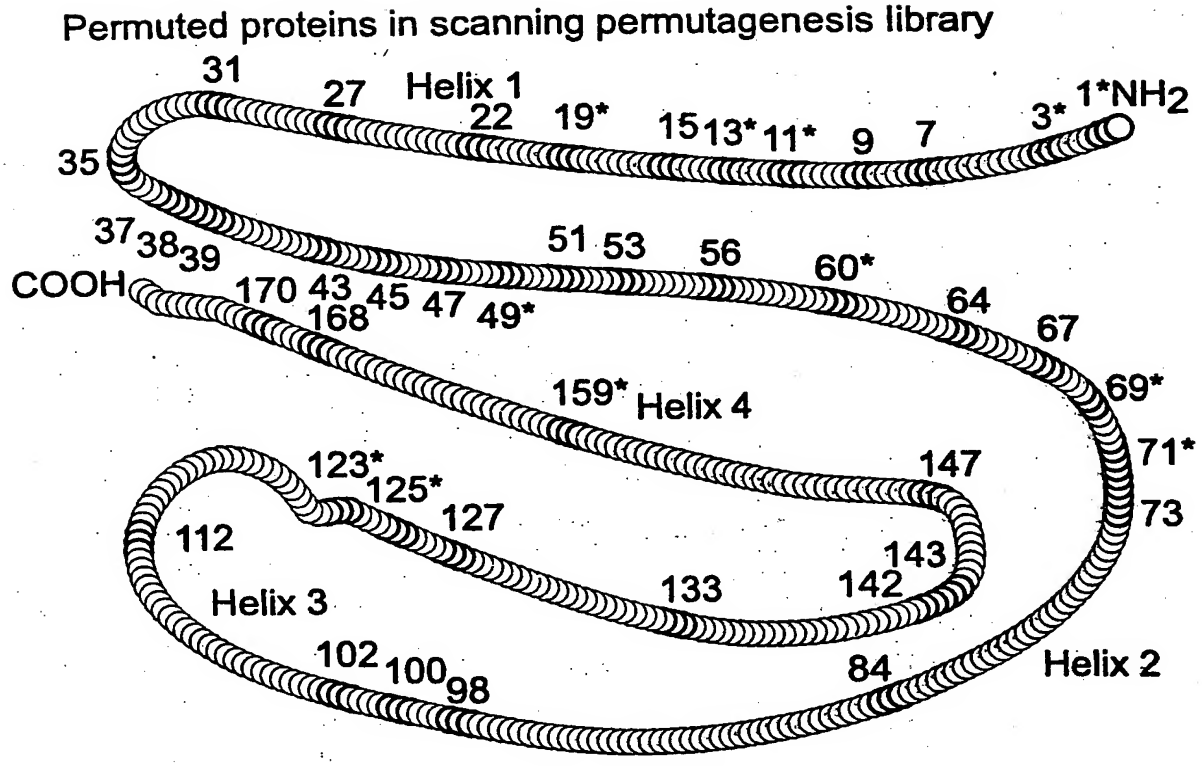
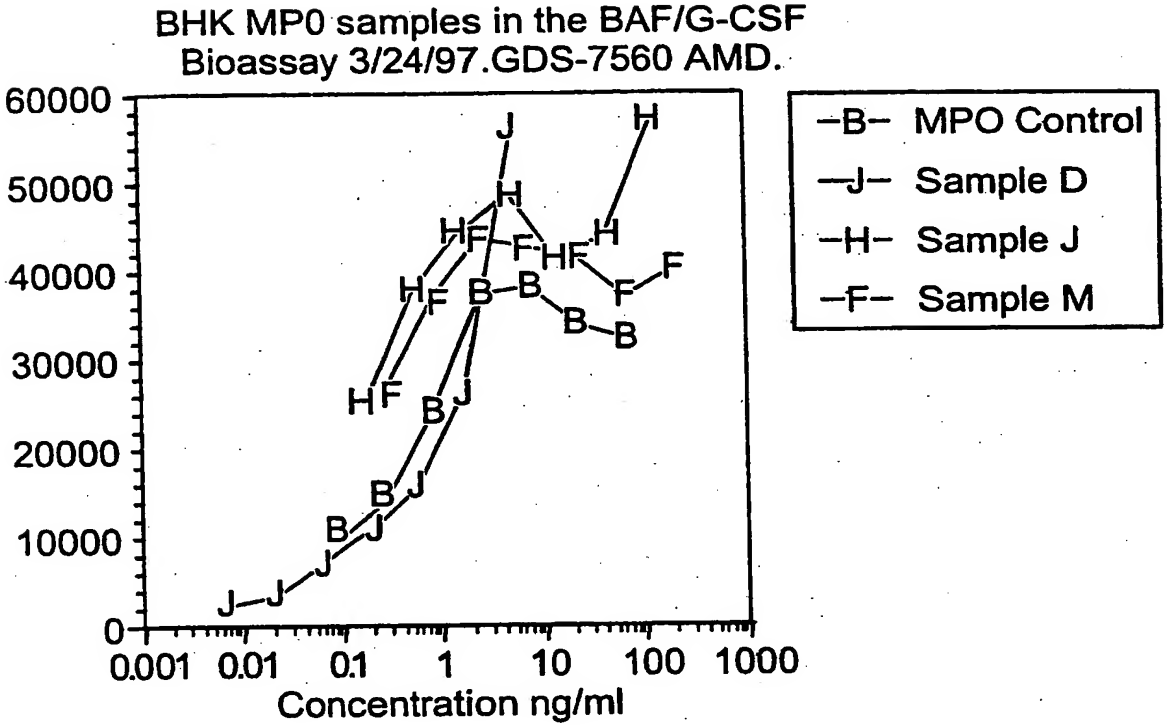

Permuted proteins in scanning permutagenesis library
31
27
Helix 1
22
19*
15
13*
11*
9
7
3*
1*NH2
35
37
38
39
51
53
56
60*
COOH
170
43
45
47
49*
168
64
67
69*
159*
Helix 4
71*
123*
125*
147
73
127
112
143
142
133
Helix 3
102
100
98
84
Helix 2

Figure 3


BHK MP0 samples in the BAF/G-CSF
Bioassay 3/24/97.GDS-7560 AMD.
60000
50000
40000
30000
20000
10000
0
0.001
0.01
0.1
1
10
100
1000
Concentration ng/ml
—B— MPO Control
—J— Sample D
—H— Sample J
—F— Sample M

SEQLIST.TXT

SEQUENCE LISTING

<110> Lee, Stephen C

<120> Method of Producing Permuteins by Scanning Permutagenesis

<130> C_3041

<140> 60/
<141> 1998-09-25

<160> 368

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 36
<212> DNA
<213> Synthetic

<400> 1
cgcgcgcaca tgtctacacc attgggccct gccagc 36

<210> 2
<211> 36
<212> DNA
<213> Synthetic

<400> 2
cgcgcgcaca tgtctccatt gggccctgcc agctcc 36

<210> 3
<211> 36
<212> DNA
<213> Synthetic

<400> 3
cgcgcgcaca tgtctttggg ccctgccagc tccctg 36

<210> 4
<211> 36
<212> DNA
<213> Synthetic

<400> 4
cgcgcgcaca tgtctggccc tgccagctcc ctgccc 36

<210> 5
<211> 36
<212> DNA
<213> Synthetic

<400> 5
cgcgcgcaca tgtctcctgc cagctccctg ccccag 36

<210> 6
<211> 36
<212> DNA
<213> Synthetic

<400> 6
cgcgcgcaca tgtctgccag ctccctgccc cagagc 36

<210> 7
<211> 36
<212> DNA
<213> Synthetic

<400> 7
cgcgcgcaca tgtctagctc cctgccccag agcttc 36

<210> 8
<211> 36
<212> DNA
<213> Synthetic

<400> 8
cgcgcgcaca tgtcttccct gccccagagc ttcctg 36

<210> 9
<211> 36
<212> DNA
<213> Synthetic

<400> 9
cgcgcgcaca tgtctctgcc ccagagcttc ctgctc 36

<210> 10
<211> 36
<212> DNA
<213> Synthetic

<400> 10
cgcgcgcaca tgtctcccca gagcttcctg ctcaag 36

<210> 11
<211> 36
<212> DNA
<213> Synthetic

<400> 11
cgcgcgcaca tgtctcagag cttcctgctc aagtct 36

<210> 12
<211> 36
<212> DNA
<213> Synthetic

<400> 12
cgcgcgcaca tgtctagctt cctgctcaag tcttta 36

<210> 13
<211> 36
<212> DNA
<213> Synthetic

<400> 13
cgcgcgcaca tgtctttcct gctcaagtct ttagag 36

<210> 14
<211> 36
<212> DNA
<213> Synthetic

<400> 14
cgcgcgcaca tgtctctgct caagtcttta gagcaa 36

<210> 15
<211> 36
<212> DNA
<213> Synthetic

<400> 15
cgcgcgcaca tgtctctcaa gtctttagag caagtg 36

<210> 16
<211> 36
<212> DNA
<213> Synthetic

<400> 16
cgcgcgcaca tgtctaagtc tttagagcaa gtgagg 36

<210> 17
<211> 36
<212> DNA
<213> Synthetic

<400> 17
cgcgcgcaca tgtcttcttt agagcaagtg aggaag 36

<210> 18
<211> 36
<212> DNA
<213> Synthetic

<400> 18
cgcgcgcaca tgtctttaga gcaagtgagg aagatc 36

<210> 19
<211> 36
<212> DNA
<213> Synthetic

<400> 19
cgcgcgcaca tgtctgagca agtgaggaag atccag 36

<210> 20
<211> 36
<212> DNA
<213> Synthetic

<400> 20
cgcgcgcaca tgtctcaagt gaggaagatc cagggc 36

<210> 21
<211> 36
<212> DNA
<213> Synthetic

<400> 21
cgcgcgcaca tgtctgtgag gaagatccag ggcgat 36

<210> 22
<211> 36
<212> DNA
<213> Synthetic

<400> 22
cgcgcgcaca tgtctaggaa gatccagggc gatggc 36

<210> 23
<211> 36
<212> DNA
<213> Synthetic

<400> 23
cgcgcgcaca tgtctaagat ccagggcgat ggcgca 36

<210> 24

<211> 36
<212> DNA
<213> Synthetic

<400> 24
cgcgcgcaca tgtctatcca gggcgatggc gcagcg 36

<210> 25
<211> 36
<212> DNA
<213> Synthetic

<400> 25
cgcgcgcaca tgtctcaggg cgatggcgca gcgctc 36

<210> 26
<211> 36
<212> DNA
<213> Synthetic

<400> 26
cgcgcgcaca tgtctggcga tggcgcagcg ctccag 36

<210> 27
<211> 36
<212> DNA
<213> Synthetic

<400> 27
cgcgcgcaca tgtctgatgg cgcagcgctc caggag 36

<210> 28
<211> 36
<212> DNA
<213> Synthetic

<400> 28
cgcgcgcaca tgtctggcgc agcgctccag gagaag 36

<210> 29
<211> 36
<212> DNA
<213> Synthetic

<400> 29
cgcgcgcaca tgtctgcagc gctccaggag aagctg 36

<210> 30
<211> 36
<212> DNA
<213> Synthetic

<400> 30
cgcgcgcaca tgtctgcgct ccaggagaag ctgtgt 36

<210> 31
<211> 36
<212> DNA
<213> Synthetic

<400> 31
cgcgcgcaca tgtctctcca ggagaagctg tgtgcc 36

<210> 32
<211> 36
<212> DNA
<213> Synthetic

<400> 32

cgcgcgcaca tgtctcagga gaagctgtgt gccacc 36

<210> 33
<211> 36
<212> DNA
<213> Synthetic

<400> 33
cgcgcgcaca tgtctgagaa gctgtgtgcc acctac 36

<210> 34
<211> 36
<212> DNA
<213> Synthetic

<400> 34
cgcgcgcaca tgtctaagct gtgtgccacc tacaag 36

<210> 35
<211> 36
<212> DNA
<213> Synthetic

<400> 35
cgcgcgcaca tgtctctgtg tgccacctac aagctg 36

<210> 36
<211> 36
<212> DNA
<213> Synthetic

<400> 36
cgcgcgcaca tgtcttgtgc cacctacaag ctgtgc 36

<210> 37
<211> 36
<212> DNA
<213> Synthetic

<400> 37
cgcgcgcaca tgtctgccac ctacaagctg tgccac 36

<210> 38
<211> 36
<212> DNA
<213> Synthetic

<400> 38
cgcgcgcaca tgtctaccta caagctgtgc cacccc 36

<210> 39
<211> 36
<212> DNA
<213> Synthetic

<400> 39
cgcgcgcaca tgtcttacaa gctgtgccac cccgag 36

<210> 40
<211> 36
<212> DNA
<213> Synthetic

<400> 40
cgcgcgcaca tgtctaagct gtgccacccc gaggag 36

<210> 41
<211> 36
<212> DNA

<213> Synthetic

<400> 41
cgcgcgcaca tgtctctgtg ccaccccgag gagctg 36

<210> 42
<211> 36
<212> DNA
<213> Synthetic

<400> 42
cgcgcgcaca tgtcttgcca ccccgaggag ctggtg 36

<210> 43
<211> 36
<212> DNA
<213> Synthetic

<400> 43
cgcgcgcaca tgtctcaccc cgaggagctg gtgctg 36

<210> 44
<211> 36
<212> DNA
<213> Synthetic

<400> 44
cgcgcgcaca tgtctcccga ggagctggtg ctgctc 36

<210> 45
<211> 36
<212> DNA
<213> Synthetic

<400> 45
cgcgcgcaca tgtctgagga gctggtgctg ctcgga 36

<210> 46
<211> 36
<212> DNA
<213> Synthetic

<400> 46
cgcgcgcaca tgtctgagct ggtgctgctc ggacac 36

<210> 47
<211> 36
<212> DNA
<213> Synthetic

<400> 47
cgcgcgcaca tgtctctggt gctgctcgga cactct 36

<210> 48
<211> 36
<212> DNA
<213> Synthetic

<400> 48
cgcgcgcaca tgtctgtgct gctcggacac tctctg 36

<210> 49
<211> 36
<212> DNA
<213> Synthetic

<400> 49
cgcgcgcaca tgtctctgct cggacactct ctgggc 36

<210> 50
<211> 36
<212> DNA
<213> Synthetic

<400> 50
cgcgcgcaca tgtctctcgg acactcctg ggcatc 36

<210> 51
<211> 36
<212> DNA
<213> Synthetic

<400> 51
cgcgcgcaca tgtctggaca ctctctgggc atcccc 36

<210> 52
<211> 36
<212> DNA
<213> Synthetic

<400> 52
cgcgcgcaca tgtctcactc tctgggcatc ccctgg 36

<210> 53
<211> 36
<212> DNA
<213> Synthetic

<400> 53
cgcgcgcaca tgtcttctct gggcatcccc tgggct 36

<210> 54
<211> 36
<212> DNA
<213> Synthetic

<400> 54
cgcgcgcaca tgtctctggg catcccctgg gctccc 36

<210> 55
<211> 36
<212> DNA
<213> Synthetic

<400> 55
cgcgcgcaca tgtctggcat cccctgggct cccctg 36

<210> 56
<211> 36
<212> DNA
<213> Synthetic

<400> 56
cgcgcgcaca tgtctatccc ctgggctccc ctgagc 36

<210> 57
<211> 36
<212> DNA
<213> Synthetic

<400> 57
cgcgcgcaca tgtctccctg ggctcccctg agctcc 36

<210> 58
<211> 36
<212> DNA
<213> Synthetic

<400> 58
cgcgcgcaca tgtcttgggc tcccctgagc tcctgc 36

<210> 59
<211> 36
<212> DNA
<213> Synthetic

<400> 59
cgcgcgcaca tgtctgctcc cctgagctcc tgcccc 36

<210> 60
<211> 36
<212> DNA
<213> Synthetic

<400> 60
cgcgcgcaca tgtctccccc gagctcctgc cccagc 36

<210> 61
<211> 36
<212> DNA
<213> Synthetic

<400> 61
cgcgcgcaca tgtctctgag ctcctgcccc agccag 36

<210> 62
<211> 36
<212> DNA
<213> Synthetic

<400> 62
cgcgcgcaca tgtctagctc ctgccccagc caggcc 36

<210> 63
<211> 36
<212> DNA
<213> Synthetic

<400> 63
cgcgcgcaca tgtcttcctg ccccagccag gccctg 36

<210> 64
<211> 36
<212> DNA
<213> Synthetic

<400> 64
cgcgcgcaca tgtcttgccc cagccaggcc ctgcag 36

<210> 65
<211> 36
<212> DNA
<213> Synthetic

<400> 65
cgcgcgcaca tgtctcccag ccaggccctg cagctg 36

<210> 66
<211> 36
<212> DNA
<213> Synthetic

<400> 66
cgcgcgcaca tgtctagcca ggccctgcag ctggca 36

<210> 67
<211> 36

<212> DNA
<213> Synthetic

<400> 67
cgcgcgcaca tgtctcaggc cctgcagctg gcaggc 36

<210> 68
<211> 36
<212> DNA
<213> Synthetic

<400> 68
cgcgcgcaca tgtctgccct gcagctggca ggctgc 36

<210> 69
<211> 36
<212> DNA
<213> Synthetic

<400> 69
cgcgcgcaca tgtctctgca gctggcaggc tgcttg 36

<210> 70
<211> 36
<212> DNA
<213> Synthetic

<400> 70
cgcgcgcaca tgtctcagct ggcaggctgc ttgagc 36

<210> 71
<211> 36
<212> DNA
<213> Synthetic

<400> 71
cgcgcgcaca tgtctctggc aggctgcttg agccaa 36

<210> 72
<211> 36
<212> DNA
<213> Synthetic

<400> 72
cgcgcgcaca tgtctgcagg ctgcttgagc caactc 36

<210> 73
<211> 36
<212> DNA
<213> Synthetic

<400> 73
cgcgcgcaca tgtctggctg cttgagccaa ctccat 36

<210> 74
<211> 36
<212> DNA
<213> Synthetic

<400> 74
cgcgcgcaca tgtcttgctt gagccaactc catagc 36

<210> 75
<211> 36
<212> DNA
<213> Synthetic

<400> 75
cgcgcgcaca tgtctttgag ccaactccat agcggc 36

<210> 76
<211> 36
<212> DNA
<213> Synthetic

<400> 76
cgcgcgcaca tgtctagcca actccatagc ggcctt 36

<210> 77
<211> 36
<212> DNA
<213> Synthetic

<400> 77
cgcgcgcaca tgtctcaact ccatagcggc cttttc 36

<210> 78
<211> 36
<212> DNA
<213> Synthetic

<400> 78
cgcgcgcaca tgtctctcca tagcggcctt ttcctc 36

<210> 79
<211> 36
<212> DNA
<213> Synthetic

<400> 79
cgcgcgcaca tgtctcatag cggccttttc ctctac 36

<210> 80
<211> 36
<212> DNA
<213> Synthetic

<400> 80
cgcgcgcaca tgtctagcgg cctttttcctc taccag 36

<210> 81
<211> 36
<212> DNA
<213> Synthetic

<400> 81
cgcgcgcaca tgtctggcct ttttctctac cagggg 36

<210> 82
<211> 36
<212> DNA
<213> Synthetic

<400> 82
cgcgcgcaca tgtctctttt cctctaccag gggctc 36

<210> 83
<211> 36
<212> DNA
<213> Synthetic

<400> 83
cgcgcgcaca tgtctttcct ctaccagggg ctcctg 36

<210> 84
<211> 36
<212> DNA
<213> Synthetic

<400> 84
cgcgcgcaca tgtctctcta ccaggggctc ctgcag 36

<210> 85
<211> 36
<212> DNA
<213> Synthetic

<400> 85
cgcgcgcaca tgtcttacca ggggctcctg caggcc 36

<210> 86
<211> 36
<212> DNA
<213> Synthetic

<400> 86
cgcgcgcaca tgtctcaggg gctcctgcag gccctg 36

<210> 87
<211> 36
<212> DNA
<213> Synthetic

<400> 87
cgcgcgcaca tgtctgggct cctgcaggcc ctggaa 36

<210> 88
<211> 36
<212> DNA
<213> Synthetic

<400> 88
cgcgcgcaca tgtctctcct gcaggccctg gaaggg 36

<210> 89
<211> 36
<212> DNA
<213> Synthetic

<400> 89
cgcgcgcaca tgtctctgca ggccctggaa gggata 36

<210> 90
<211> 36
<212> DNA
<213> Synthetic

<400> 90
cgcgcgcaca tgtctcaggc cctggaaggg atatcc 36

<210> 91
<211> 36
<212> DNA
<213> Synthetic

<400> 91
cgcgcgcaca tgtctgccct ggaagggata tccccc 36

<210> 92
<211> 36
<212> DNA
<213> Synthetic

<400> 92
cgcgcgcaca tgtctctgga agggatatcc cccgag 36

<210> 93

<211> 36
<212> DNA
<213> Synthetic

<400> 93
cgcgcgcaca tgtctgaagg gatatccccc gagttg 36

<210> 94
<211> 36
<212> DNA
<213> Synthetic

<400> 94
cgcgcgcaca tgtctgggat atcccccgag ttgggt 36

<210> 95
<211> 36
<212> DNA
<213> Synthetic

<400> 95
cgcgcgcaca tgtctatatc ccccgagttg ggtccc 36

<210> 96
<211> 36
<212> DNA
<213> Synthetic

<400> 96
cgcgcgcaca tgtcttcccc cgagttgggt cccacc 36

<210> 97
<211> 36
<212> DNA
<213> Synthetic

<400> 97
cgcgcgcaca tgtctcccga gttgggtccc accttg 36

<210> 98
<211> 36
<212> DNA
<213> Synthetic

<400> 98
cgcgcgcaca tgtctgagtt gggtcccacc ttggac 36

<210> 99
<211> 36
<212> DNA
<213> Synthetic

<400> 99
cgcgcgcaca tgtctttggg tcccaccttg gacaca 36

<210> 100
<211> 36
<212> DNA
<213> Synthetic

<400> 100
cgcgcgcaca tgtctggtcc caccttggac acactg 36

<210> 101
<211> 36
<212> DNA
<213> Synthetic

<400> 101

cgcgcgcaca tgtctcccac cttggacaca ctgcag 36

<210> 102
<211> 36
<212> DNA
<213> Synthetic

<400> 102
cgcgcgcaca tgtctacctt ggacacactg cagctg 36

<210> 103
<211> 36
<212> DNA
<213> Synthetic

<400> 103
cgcgcgcaca tgtctttgga cacactgcag ctggac 36

<210> 104
<211> 36
<212> DNA
<213> Synthetic

<400> 104
cgcgcgcaca tgtctgacac actgcagctg gacgtc 36

<210> 105
<211> 36
<212> DNA
<213> Synthetic

<400> 105
cgcgcgcaca tgtctacact gcagctggac gtcgcc 36

<210> 106
<211> 36
<212> DNA
<213> Synthetic

<400> 106
cgcgcgcaca tgtctctgca gctggacgtc gccgac 36

<210> 107
<211> 36
<212> DNA
<213> Synthetic

<400> 107
cgcgcgcaca tgtctcagct ggacgtcgcc gacttt 36

<210> 108
<211> 36
<212> DNA
<213> Synthetic

<400> 108
cgcgcgcaca tgtctctgga cgtcgccgac tttgcc 36

<210> 109
<211> 36
<212> DNA
<213> Synthetic

<400> 109
cgcgcgcaca tgtctgacgt cgccgacttt gccacc 36

<210> 110
<211> 36
<212> DNA

<213> Synthetic

<400> 110
cgcgcgcaca tgtctgtcgc cgactttgcc accacc 36

<210> 111
<211> 36
<212> DNA
<213> Synthetic

<400> 111
cgcgcgcaca tgtctgccga ctttgccacc accatc 36

<210> 112
<211> 36
<212> DNA
<213> Synthetic

<400> 112
cgcgcgcaca tgtctgactt tgccaccacc atctgg 36

<210> 113
<211> 36
<212> DNA
<213> Synthetic

<400> 113
cgcgcgcaca tgtcttttgc caccaccatc tggcag 36

<210> 114
<211> 36
<212> DNA
<213> Synthetic

<400> 114
cgcgcgcaca tgtctgccac caccatctgg cagcag 36

<210> 115
<211> 36
<212> DNA
<213> Synthetic

<400> 115
cgcgcgcaca tgtctaccac catctggcag cagatg 36

<210> 116
<211> 36
<212> DNA
<213> Synthetic

<400> 116
cgcgcgcaca tgtctaccat ctggcagcag atggaa 36

<210> 117
<211> 36
<212> DNA
<213> Synthetic

<400> 117
cgcgcgcaca tgtctatctg gcagcagatg gaagaa 36

<210> 118
<211> 36
<212> DNA
<213> Synthetic

<400> 118
cgcgcgcaca tgtcttggca gcagatggaa gaactg 36

<210> 119
<211> 36
<212> DNA
<213> Synthetic

<400> 119
cgcgcgcaca tgtctcagca gatggaagaa ctggga 36

<210> 120
<211> 36
<212> DNA
<213> Synthetic

<400> 120
cgcgcgcaca tgtctcagat ggaagaactg ggaatg 36

<210> 121
<211> 36
<212> DNA
<213> Synthetic

<400> 121
cgcgcgcaca tgtctatgga agaactggga atggcc 36

<210> 122
<211> 36
<212> DNA
<213> Synthetic

<400> 122
cgcgcgcaca tgtctgaaga actgggaatg gcccct 36

<210> 123
<211> 36
<212> DNA
<213> Synthetic

<400> 123
cgcgcgcaca tgtctgaact gggaatggcc cctgcc 36

<210> 124
<211> 36
<212> DNA
<213> Synthetic

<400> 124
cgcgcgcaca tgtctctggg aatggccccc gccctg 36

<210> 125
<211> 36
<212> DNA
<213> Synthetic

<400> 125
cgcgcgcaca tgtctggaat ggcccctgcc ctgcag 36

<210> 126
<211> 36
<212> DNA
<213> Synthetic

<400> 126
cgcgcgcaca tgtctatggc ccctgccctg cagccc 36

<210> 127
<211> 36
<212> DNA
<213> Synthetic

<400> 127
cgcgcgcaca tgtctgcccc tgccctgcag cccacc 36

<210> 128
<211> 36
<212> DNA
<213> Synthetic

<400> 128
cgcgcgcaca tgtctcctgc cctgcagccc acccag 36

<210> 129
<211> 36
<212> DNA
<213> Synthetic

<400> 129
cgcgcgcaca tgtctgccct gcagcccacc cagggt 36

<210> 130
<211> 36
<212> DNA
<213> Synthetic

<400> 130
cgcgcgcaca tgtctctgca gcccacccag ggtgcc 36

<210> 131
<211> 36
<212> DNA
<213> Synthetic

<400> 131
cgcgcgcaca tgtctcagcc cacccagggt gccatg 36

<210> 132
<211> 36
<212> DNA
<213> Synthetic

<400> 132
cgcgcgcaca tgtctcccac ccagggtgcc atgccg 36

<210> 133
<211> 36
<212> DNA
<213> Synthetic

<400> 133
cgcgcgcaca tgtctaccca gggtgccatg ccggcc 36

<210> 134
<211> 36
<212> DNA
<213> Synthetic

<400> 134
cgcgcgcaca tgtctcaggg tgccatgccg gccttc 36

<210> 135
<211> 36
<212> DNA
<213> Synthetic

<400> 135
cgcgcgcaca tgtctggtgc catgccggcc ttcgcc 36

<210> 136
<211> 36

<212> DNA
<213> Synthetic

<400> 136
cgcgcgcaca tgtctgccat gccggccttc gcctct 36

<210> 137
<211> 36
<212> DNA
<213> Synthetic

<400> 137
cgcgcgcaca tgtctatgcc ggccttcgcc tctgct 36

<210> 138
<211> 36
<212> DNA
<213> Synthetic

<400> 138
cgcgcgcaca tgtctccggc cttcgcctct gctttc 36

<210> 139
<211> 36
<212> DNA
<213> Synthetic

<400> 139
cgcgcgcaca tgtctgcctt cgcctctgct ttccag 36

<210> 140
<211> 36
<212> DNA
<213> Synthetic

<400> 140
cgcgcgcaca tgtctttcgc ctctgctttc cagcgc 36

<210> 141
<211> 36
<212> DNA
<213> Synthetic

<400> 141
cgcgcgcaca tgtctgcctc tgctttccag cgccgg 36

<210> 142
<211> 36
<212> DNA
<213> Synthetic

<400> 142
cgcgcgcaca tgtcttctgc tttccagcgc cgggca 36

<210> 143
<211> 36
<212> DNA
<213> Synthetic

<400> 143
cgcgcgcaca tgtctgcttt ccagcgccgg gcagga 36

<210> 144
<211> 36
<212> DNA
<213> Synthetic

<400> 144
cgcgcgcaca tgtctttcca gcgccgggca ggaggg 36

<210> 145
<211> 36
<212> DNA
<213> Synthetic

<400> 145
cgcgcgcaca tgtctcagcg ccgggcagga ggggtc 36

<210> 146
<211> 36
<212> DNA
<213> Synthetic

<400> 146
cgcgcgcaca tgtctcgccg ggcaggaggg gtcctg 36

<210> 147
<211> 36
<212> DNA
<213> Synthetic

<400> 147
cgcgcgcaca tgtctcgggc aggaggggtc ctggtt 36

<210> 148
<211> 36
<212> DNA
<213> Synthetic

<400> 148
cgcgcgcaca tgtctgcagg aggggtcctg gttgct 36

<210> 149
<211> 36
<212> DNA
<213> Synthetic

<400> 149
cgcgcgcaca tgtctggagg ggtcctggtt gctagc 36

<210> 150
<211> 36
<212> DNA
<213> Synthetic

<400> 150
cgcgcgcaca tgtctggggt cctggttgct agccat 36

<210> 151
<211> 36
<212> DNA
<213> Synthetic

<400> 151
cgcgcgcaca tgtctgtcct ggttgctagc catctg 36

<210> 152
<211> 36
<212> DNA
<213> Synthetic

<400> 152
cgcgcgcaca tgtctctggt tgctagccat ctgcag 36

<210> 153
<211> 36
<212> DNA
<213> Synthetic

<400> 153
cgcgcgcaca tgtctgttgc tagccatctg cagagc 36

<210> 154
<211> 36
<212> DNA
<213> Synthetic

<400> 154
cgcgcgcaca tgtctgctag ccatctgcag agcttc 36

<210> 155
<211> 36
<212> DNA
<213> Synthetic

<400> 155
cgcgcgcaca tgtctagcca tctgcagagc ttcctg 36

<210> 156
<211> 36
<212> DNA
<213> Synthetic

<400> 156
cgcgcgcaca tgtctcatct gcagagcttc ctggag 36

<210> 157
<211> 36
<212> DNA
<213> Synthetic

<400> 157
cgcgcgcaca tgtctctgca gagcttcctg gaggtg 36

<210> 158
<211> 36
<212> DNA
<213> Synthetic

<400> 158
cgcgcgcaca tgtctcagag cttcctggag gtgtcg 36

<210> 159
<211> 36
<212> DNA
<213> Synthetic

<400> 159
cgcgcgcaca tgtctagctt cctggaggtg tcgtac 36

<210> 160
<211> 36
<212> DNA
<213> Synthetic

<400> 160
cgcgcgcaca tgtctttcct ggaggtgtcg taccgc 36

<210> 161
<211> 36
<212> DNA
<213> Synthetic

<400> 161
cgcgcgcaca tgtctctgga ggtgtcgtac cgcgtt 36

<210> 162

<211> 36
<212> DNA
<213> Synthetic

<400> 162
cgcgcgcaca tgtctgaggt gtcgtaccgc gttcta 36

<210> 163
<211> 36
<212> DNA
<213> Synthetic

<400> 163
cgcgcgcaca tgtctgtgtc gtaccgcgtt ctacgc 36

<210> 164
<211> 36
<212> DNA
<213> Synthetic

<400> 164
cgcgcgcaca tgtcttcgta ccgcgttcta cgccac 36

<210> 165
<211> 36
<212> DNA
<213> Synthetic

<400> 165
cgcgcgcaca tgtcttaccg cgttctacgc caccttt 36

<210> 166
<211> 36
<212> DNA
<213> Synthetic

<400> 166
cgcgcgcaca tgtctcgcgt tctacgccac cttgcg 36

<210> 167
<211> 36
<212> DNA
<213> Synthetic

<400> 167
cgcgcgcaca tgtctgttct acgccacctt gcgcag 36

<210> 168
<211> 36
<212> DNA
<213> Synthetic

<400> 168
cgcgcgcaca tgtctctacg ccaccttgcg cagccc 36

<210> 169
<211> 36
<212> DNA
<213> Synthetic

<400> 169
cgcgcgcaca tgtctcgcca ccttgcgcag cccgac 36

<210> 170
<211> 36
<212> DNA
<213> Synthetic

<400> 170

cgcgcgcaca tgtctcacct tgcgcagccc gacatg 36

<210> 171
<211> 36
<212> DNA
<213> Synthetic

<400> 171
cgcgcgcaca tgtctcttgc gcagcccgac atggct 36

<210> 172
<211> 36
<212> DNA
<213> Synthetic

<400> 172
cgcgcgcaca tgtctgcgca gcccgacatg gctaca 36

<210> 173
<211> 36
<212> DNA
<213> Synthetic

<400> 173
cgcgcgcaca tgtctcagcc cgacatggct acacca 36

<210> 174
<211> 36
<212> DNA
<213> Synthetic

<400> 174
cgcgcgcaca tgtctcccga catggctaca ccattg 36

<210> 175
<211> 36
<212> DNA
<213> Synthetic

<400> 175
cgcgcgcaca tgtctcgcca ccttgcgcag cccact 36

<210> 176
<211> 36
<212> DNA
<213> Synthetic

<400> 176
cgcgcgcaca tgtctcacct tgcgcagccc actagt 36

<210> 177
<211> 36
<212> DNA
<213> Synthetic

<400> 177
cgcgcgcaca tgtctcttgc gcagcccact agtcat 36

<210> 178
<211> 36
<212> DNA
<213> Synthetic

<400> 178
cgcgcgcaca tgtctgcgca gcccactagt catcca 36

<210> 179
<211> 36
<212> DNA

<213> Synthetic

<400> 179
cgcgcgcaca tgtctcagcc cactagtcat ccacct 36

<210> 180
<211> 36
<212> DNA
<213> Synthetic

<400> 180
cgcgcgcaca tgtctcccac tagtcatcca cctatg 36

<210> 181
<211> 36
<212> DNA
<213> Synthetic

<400> 181
cgcgcgcaca tgtctcgcca ccttgcgcag cccggc 36

<210> 182
<211> 36
<212> DNA
<213> Synthetic

<400> 182
cgcgcgcaca tgtctcacct tgcgcagccc ggcggc 36

<210> 183
<211> 36
<212> DNA
<213> Synthetic

<400> 183
cgcgcgcaca tgtctcttgc gcagcccggc ggcggc 36

<210> 184
<211> 36
<212> DNA
<213> Synthetic

<400> 184
cgcgcgcaca tgtctgcgca gcccggcggc ggctct 36

<210> 185
<211> 36
<212> DNA
<213> Synthetic

<400> 185
cgcgcgcaca tgtctcagcc cggcggcggc tctgac 36

<210> 186
<211> 36
<212> DNA
<213> Synthetic

<400> 186
cgcgcgcaca tgtctcccgg cggcggctct gacatg 36

<210> 187
<211> 36
<212> DNA
<213> Synthetic

<400> 187
tatatatgcg gccgcagcca tgtcgggctg cgcaag 36

<210> 188
<211> 36
<212> DNA
<213> Synthetic

<400> 188
tatatatgcg gccgcagcca tgtcacgcgt acgatt 36

<210> 189
<211> 36
<212> DNA
<213> Synthetic

<400> 189
tatatatgcg gccgcagcca tgtcagagcc gccgcc 36

<210> 190
<211> 36
<212> DNA
<213> Synthetic

<400> 190
tatatatgcg gccgctgtag ccatgtcggg ctgcgc 36

<210> 191
<211> 36
<212> DNA
<213> Synthetic

<400> 191
tatatatgcg gccgctgtag ccatgtcacg cgtacg 36

<210> 192
<211> 36
<212> DNA
<213> Synthetic

<400> 192
tatatatgcg gccgctgtag ccatgtcaga gccgcc 36

<210> 193
<211> 36
<212> DNA
<213> Synthetic

<400> 193
tatatatgcg gccgctggtg tagccatgtc gggctg 36

<210> 194
<211> 36
<212> DNA
<213> Synthetic

<400> 194
tatatatgcg gccgctggtg tagccatgtc acgcgt 36

<210> 195
<211> 36
<212> DNA
<213> Synthetic

<400> 195
tatatatgcg gccgctggtg tagccatgtc agagcc 36

<210> 196
<211> 36
<212> DNA
<213> Synthetic

<400> 196
tatatatgcg gccgccaatg gtgtagccat gtcggg 36

<210> 197
<211> 36
<212> DNA
<213> Synthetic

<400> 197
tatatatgcg gccgccaatg gtgtagccat gtcacg 36

<210> 198
<211> 36
<212> DNA
<213> Synthetic

<400> 198
tatatatgcg gccgccaatg gtgtagccat gtcaga 36

<210> 199
<211> 36
<212> DNA
<213> Synthetic

<400> 199
tatatatgcg gccgcgccca atggtgtagc catgtc 36

<210> 200
<211> 36
<212> DNA
<213> Synthetic

<400> 200
tatatatgcg gccgcagggc ccaatggtgt agccat 36

<210> 201
<211> 36
<212> DNA
<213> Synthetic

<400> 201
tatatatgcg gccgcggcag ggcccaatgg tgtagc 36

<210> 202
<211> 36
<212> DNA
<213> Synthetic

<400> 202
tatatatgcg gccgcgctgg cagggcccaa tggtgt 36

<210> 203
<211> 36
<212> DNA
<213> Synthetic

<400> 203
tatatatgcg gccgcggagc tggcagggcc caatgg 36

<210> 204
<211> 36
<212> DNA
<213> Synthetic

<400> 204
tatatatgcg gccgcccagg gctggccagg gcccaa 36

<210> 205
<211> 36

<212> DNA
<213> Synthetic

<400> 205
tatatatgcg gccgcgggca gggagctggc agggcc 36

<210> 206
<211> 36
<212> DNA
<213> Synthetic

<400> 206
tatatatgcg gccgcctggg gcagggagct ggcagg 36

<210> 207
<211> 36
<212> DNA
<213> Synthetic

<400> 207
tatatatgcg gccgcgctct ggggcaggga gctggc 36

<210> 208
<211> 36
<212> DNA
<213> Synthetic

<400> 208
tatatatgcg gccgcgaagc tctggggcag ggagct 36

<210> 209
<211> 36
<212> DNA
<213> Synthetic

<400> 209
tatatatgcg gccgccagga agctctgggg caggga 36

<210> 210
<211> 36
<212> DNA
<213> Synthetic

<400> 210
tatatatgcg gccgcgagca ggaagctctg gggcag 36

<210> 211
<211> 36
<212> DNA
<213> Synthetic

<400> 211
tatatatgcg gccgccttga gcaggaagct ctgggg 36

<210> 212
<211> 36
<212> DNA
<213> Synthetic

<400> 212
tatatatgcg gccgcagact tgagcaggaa gctctg 36

<210> 213
<211> 36
<212> DNA
<213> Synthetic

<400> 213
tatatatgcg gccgctaaag acttgagcag gaagct 36

<210> 214
<211> 36
<212> DNA
<213> Synthetic

<400> 214
tatatatgcg gccgcctcta aagacttgag caggaa 36

<210> 215
<211> 36
<212> DNA
<213> Synthetic

<400> 215
tatatatgcg gccgcttgct ctaaagactt gagcag 36

<210> 216
<211> 36
<212> DNA
<213> Synthetic

<400> 216
tatatatgcg gccgccactt gctctaaaga cttgag 36

<210> 217
<211> 36
<212> DNA
<213> Synthetic

<400> 217
tatatatgcg gccgccctca cttgctctaa agactt 36

<210> 218
<211> 36
<212> DNA
<213> Synthetic

<400> 218
tatatatgcg gccgccttcc tcacttgctc taaaga 36

<210> 219
<211> 36
<212> DNA
<213> Synthetic

<400> 219
tatatatgcg gccgcgatct tcctcacttg ctctaa 36

<210> 220
<211> 36
<212> DNA
<213> Synthetic

<400> 220
tatatatgcg gccgcctgga tcttcctcac ttgctc 36

<210> 221
<211> 36
<212> DNA
<213> Synthetic

<400> 221
tatatatgcg gccgcgccct ggatcttcct cacttg 36

<210> 222
<211> 36
<212> DNA
<213> Synthetic

<400> 222
tatatatgcg gccgcatcgc cctggatctt cctcac 36

<210> 223
<211> 36
<212> DNA
<213> Synthetic

<400> 223
tatatatgcg gccgcgccat cgccctggat cttcct 36

<210> 224
<211> 36
<212> DNA
<213> Synthetic

<400> 224
tatatatgcg gccgctgcgc catcgccctg gatctt 36

<210> 225
<211> 36
<212> DNA
<213> Synthetic

<400> 225
tatatatgcg gccgccgctg cgccatcgcc ctggat 36

<210> 226
<211> 36
<212> DNA
<213> Synthetic

<400> 226
tatatatgcg gccgcgagcg ctgcgccatc gccctg 36

<210> 227
<211> 36
<212> DNA
<213> Synthetic

<400> 227
tatatatgcg gccgcctgga gcgctgcgcc atcgcc 36

<210> 228
<211> 36
<212> DNA
<213> Synthetic

<400> 228
tatatatgcg gccgcctcct ggagcgctgc gccatc 36

<210> 229
<211> 36
<212> DNA
<213> Synthetic

<400> 229
tatatatgcg gccgccttct cctggagcgc tgcgcc 36

<210> 230
<211> 36
<212> DNA
<213> Synthetic

<400> 230
tatatatgcg gccgccagct tctcctggag cgctgc 36

<210> 231